

Pending Nucleic Acid and Pending Amino Acid database searches generate two sets of results each. The Pending databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Searches run against the Nucleic Acid Pending database produce two sets of results, with the extensions **.rnpm** and **.rnpn**

Searches run against the Amino Acid Pending database produce two sets of results, with the extensions **.rapm** and **.rapn**

Because they contain data that is confidential, the results of Pending database searches should not be left in the case .

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 14, 2004, 08:23:58 ; Search time 60 Seconds
(without alignments)
1323.264 Million cell updates/sec

Title: US-09-931-232-1

Perfect score: 1 MWSKGEBLFTGVPELVELD.....AQESGMDRHPACASARINV 281

Sequence:

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1500	100.0	281	3	AAV50142
2	1500	100.0	281	3	AAV50142
3	1500	100.0	281	3	AAV50142
4	1289.5	86.0	805	3	AAU10888
5	1289.5	86.0	805	3	AAU10888
6	1280.5	85.4	793	5	ABG94503
7	1280.5	85.4	793	5	ABG94503
8	1279.5	85.3	357	6	ABP72573
9	1279.5	85.3	357	6	ABP72573
10	1278	85.2	607	2	AAW85032
11	1278	85.2	607	2	AAW85032
12	1277	85.1	1089	7	ADH90657
13	1276.5	85.1	284	7	ADH90657
14	1276.5	85.1	1090	7	ADH90657
15	1276	85.1	286	7	ADH90657
16	1276	85.1	289	7	ADH90657
17	1276	85.1	290	7	ADH90657
18	1276	85.1	290	7	ADH90657
19	1276	85.1	290	7	ADH90657
20	1276	85.1	290	7	ADH90657
21	1276	85.1	290	7	ADH90657
22	1276	85.1	290	7	ADH90657
23	1275.5	85.0	941	3	AAW85039
24	1275.5	85.0	941	3	AAW85039
25	1275.5	85.0	941	3	AAW85039

26	1275	85.0	432	6	ABP56093
27	1275	85.0	727	2	AAW85041
28	1275	85.0	797	2	AAW85043
29	1275	85.0	797	3	AAV70782
30	1275	85.0	918	3	AAV70782
31	1274.5	85.0	255	4	AAW86143
32	1274	84.9	239	3	AAW86143
33	1274	84.9	239	3	AAW86143
34	1274	84.9	239	3	AAW86143
35	1274	84.9	239	3	AAW86143
36	1274	84.9	239	3	AAW86143
37	1274	84.9	239	3	AAW86143
38	1274	84.9	239	3	AAW86143
39	1274	84.9	239	3	AAW86143
40	1274	84.9	239	3	AAW86143
41	1274	84.9	239	3	AAW86143
42	1274	84.9	239	3	AAW86143
43	1274	84.9	239	3	AAW86143
44	1274	84.9	239	3	AAW86143
45	1274	84.9	239	3	AAW86143

ALIGNMENTS

RESULT 1

AAV50142 standard; protein; 281 AA.

AAV50142;

12-SEP-2003 (revised)

31-JAN-2000 (first entry)

Green fluorescent protein/ornithine decarboxylase fusion gene.

Enhanced green fluorescent protein; EGFP; ornithine decarboxylase; MODC; fusion; reporter; degradation domain; C-terminal; gene expression; protein localization; fluorescence; short-lived; destabilized; half-life; rapid degradation; EST sequence; EGFP-MODC; EGFP-MODC422-461; rapid turnover; determination; short-term event; repetitive event; sensitive; developmental process; transient; periodic; cyclic; expression; protein transport; circadian rhythm; kinetic; dynamic; transcription.

Aequorea victoria.

Mus sp.

Chimeric.

WC9954348-A1.

28-OCT-1999.

13-NOV-1998; 98WO-US024323.

17-APR-1998; 98US-00062102.

(CLON-) CLONTECH LAB INC.

Li X, Kain S;

WPI; 2000-013225/01.

N-PSDB; AA232767.

New fusion protein useful for studying cell localization, for studying cell lineage and for assaying activation or deactivation of transcriptional or translational elements.

Claim 5; Page 23; 49p; English.

This sequence represents a novel fusion protein, EGFP-MODC422-461, comprising an enhanced variant of green fluorescent protein (EGFP) fused via its C-terminus to residues 422-461 of the C-terminal degradation

CC domain of mouse ornithine decarboxylase (MODC). Green fluorescent protein (GFP) is a widely used reporter in studies of gene expression and protein localization; however, it is a very stable protein which can therefore accumulate and thus is often toxic to mammalian cells. MODC, in contrast, is on of the most short-lived proteins in mammalian cells, with a half life of about 30 minutes. Its rapid degradation is due to the composition of its C-terminus, a portion of which contains a PST sequence. The PST sequence contains a region enriched with proline, glutamate, serine and threonine, often flanked by basic amino acids, and has been proposed as characterizing short-lived proteins. The EGFP-MODC fusion protein, which comprises EGFP and the C-terminal MODC degradation domain, including a PST sequence, combines the fluorescence of GFP with the short life of MODC, and has a fluorescence half-life of about 2 hours in vivo. As the new fusion protein is destabilised with a rapid turnover, it is useful for producing stable cell lines which express GFP without the toxicity caused by build-up of GFP. Unlike prior art GFPs, the new fusion protein is useful for determining short-term or repetitive events, and produces more sensitive results due to the lack of accumulation of GFP. GFP fluorescence can be detected without additional expensive steps, e.g., lysing cells, adding exogenous substrates. The destabilised EGFP is especially useful for studying developmental processes where genes are transiently expressed, dynamics of protein transport, and periodic and cyclic expression of genes that control unique biological phenomena such as circadian rhythms, and is useful as a reporter to study the kinetics of mRNA transcription from a regulated promoter. (Updated on 12-SEP-2003 to standardise OS field)

Sequence 281 AA;

Query Match 100.0%; Score 1500; DB 3; Length 281;
Best Local Similarity 100.0%; Pred. No. 6.7e-146;
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MWSKBEHFTGVVPLIVELDGVNCHKPSVSGEGDATTGKLTLCFTCTGKLPVWPPT 60
DB 1 MWSKBEHFTGVVPLIVELDGVNCHKPSVSGEGDATTGKLTLCFTCTGKLPVWPPT 60
QY 61 LVTLTYGVQCFSRYPDMKHQDFPKSAMPEGYQERTIFFKDDGNKYKTRAEVFEGBPTL 120
DB 61 LVTLTYGVQCFSRYPDMKHQDFPKSAMPEGYQERTIFFKDDGNKYKTRAEVFEGBPTL 120
QY 121 VNRLEIKGIDPEKEDNIIIGHKLEYNVNSHNYIMADKQNGIKVNFKIRHNIEDGSVOLA 180
DB 121 VNRLEIKGIDPEKEDNIIIGHKLEYNVNSHNYIMADKQNGIKVNFKIRHNIEDGSVOLA 180
QY 181 DHYQONTPIGPGPVLDPDNHYLSTOSALSCKDPNEKRDMVLLFEVTAAGITLGMDELTKK 240
DB 181 DHYQONTPIGPGPVLDPDNHYLSTOSALSCKDPNEKRDMVLLFEVTAAGITLGMDELTKK 240
QY 241 LSHGFPPEVEBODDGTLPMSCAQESGMDRHPAACASARINV 281
DB 241 LSHGFPPEVEBODDGTLPMSCAQESGMDRHPAACASARINV 281

RESULT 2

AAB24252
ID AAB24252 standard; protein; 281 AA.

AC AAB24252;

DT 08-FEB-2001 (first entry)

DE EGFP-MODC422-461 fusion protein SEQ ID NO:1.

KM Green fluorescent protein; GFP; jellyfish; EGFP; mouse; ODC; MODC;
KW ornithine decarboxylase; PST; fusion protein; genetic reporter;
XX identification.

OS Aequorea victoria.

OS Mus musculus.

OS Synthetic.

PN US6130313-A.

XX 10-OCT-2000.
XX
XX
PF 17-APR-1998; 98US-00062102.
XX
XX 02-OCT-1997; 97US-0060855P.
XX
XX (CLON-) CLONTECH LAB INC.
XX
XX Kain S, Li X;
XX
XX WPI; 2000-013225/01.
DR N-PSDB; AAC64428.
XX
XX New fusion protein useful for studying cell localization, for studying
PT cell lineage and for assaying activation or deactivation of
PT transcriptional or translational elements.
XX
XX Claim 5; Col 12; 18pp; English.

Sequence 281 AA;

Query Match 100.0%; Score 1500; DB 3; Length 281;
Best Local Similarity 100.0%; Pred. No. 6.7e-146;
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MWSKBEHFTGVVPLIVELDGVNCHKPSVSGEGDATTGKLTLCFTCTGKLPVWPPT 60
DB 1 MWSKBEHFTGVVPLIVELDGVNCHKPSVSGEGDATTGKLTLCFTCTGKLPVWPPT 60
QY 61 LVTLTYGVQCFSRYPDMKHQDFPKSAMPEGYQERTIFFKDDGNKYKTRAEVFEGBPTL 120
DB 61 LVTLTYGVQCFSRYPDMKHQDFPKSAMPEGYQERTIFFKDDGNKYKTRAEVFEGBPTL 120
QY 121 VNRLEIKGIDPEKEDNIIIGHKLEYNVNSHNYIMADKQNGIKVNFKIRHNIEDGSVOLA 180
DB 121 VNRLEIKGIDPEKEDNIIIGHKLEYNVNSHNYIMADKQNGIKVNFKIRHNIEDGSVOLA 180
QY 181 DHYQONTPIGPGPVLDPDNHYLSTOSALSCKDPNEKRDMVLLFEVTAAGITLGMDELTKK 240
DB 181 DHYQONTPIGPGPVLDPDNHYLSTOSALSCKDPNEKRDMVLLFEVTAAGITLGMDELTKK 240
QY 241 LSHGFPPEVEBODDGTLPMSCAQESGMDRHPAACASARINV 281
DB 241 LSHGFPPEVEBODDGTLPMSCAQESGMDRHPAACASARINV 281

RESULT 3

AAU10888
ID AAU10888 standard; protein; 281 AA.

AC AAU10888;

DT 29-AUG-2003 (revised)

DT 14-FEB-2002 (first entry)

DE EGFP-MODC422-461 fusion protein.

KM Green fluorescent protein; ornithine decarboxylase; EGFP/MODC422-461;
KW PST region; protein degradation; drug screening;
XX transcriptional control sequence; translational control sequence.

OS Aequorea victoria.

OS Mus sp.

OS Chimeric.

```

XX Key Location/Qualifiers
FH 1..242
FT Protein /label= Green fluorescent protein
FT Peptide 243..281
FT /label= MODC422_461
FT /note= "Murine ornithine decarboxylase residues 442-461"
FT 243..270
FT /label= PE8T_region
FT /note= "Region enriched with Pro, Glu, Ser and Thr,
FT proposed to confer a short half-life on the protein"
XX
XX US6306600-B1.
XX
XX 23-OCT-2001.
XX
XX 30-JUL-1999; 99US-00364946.
XX
XX 17-APR-1998; 98US-00062102.
XX 13-NOV-1998; 98US-00191233.
XX
XX (CLON-) CLONTECH LAB INC.
XX
XX Kain S, Li X;
XX
XX MPI; 2002-040198/05.
XX N-PSDB; AAS16105.
XX
XX Assaying the regulatory function of a transcriptional control sequence,
XX for drug screening or for analyzing protein degradation, comprises
XX employing rapidly degrading Green Fluorescent Protein-fusion proteins.
XX
XX Example 6; Col 12; 25pp; English.
XX
XX The invention relates to assaying the regulatory function of a
XX transcriptional or translational control sequence, with a transient
XX fluorescent reporter protein, comprises transfecting cells with an
XX expression vector having a DNA sequence that codes for a fluorescent
XX fusion protein comprising a Green Fluorescent Protein (GFP) and a PE8T (a
XX region enriched for Pro, Glu, Ser and Thr residues) sequence. The fusion
XX protein, due to the presence of the PE8T sequence has a half-life of no
XX ore than 10 hours. The method is useful for analysing transcriptional
XX regulation and/or cis-acting regulatory elements. The method is also
XX useful for studying protein degradation. The method is also useful for
XX studying processes involving multiple gene expression, as well as drug
XX screening assays. The present method permits easier development of stable
XX cell lines that express the GFP gene, since toxic levels of GFP are
XX avoided because the GFP protein is degraded quickly. Furthermore, the
XX destabilised EGFP-MODC (a humanised GFP/murine ornithine decarboxylase
XX PE8T region fusion protein) decreases GFP accumulation. Accumulation of
XX fluorescent protein can interfere with the sensitivity of analysis. Thus,
XX the destabilised, rapid turnover fusion protein renders more sensitive
XX results. The present sequence is EGFP-MODC422-461, a humanised GFP/murine
XX ornithine decarboxylase PE8T region fusion protein for use in the method
XX of the invention. (Updated on 29-AUG-2003 to standardise OS field)
XX
XX Sequence 281 AA;
XX
XX Query Match 100.0%; Score 1500; DB 5; Length 281;
XX Best Local Similarity 100.0%; Pred. No. 6,7e-146;
XX Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 MWSKGEELFTGVPLFLVELDGVNCHKFSVSGEGDATTYGLTKLFTCTTGKLPVPMPT 60
XX 1 MWSKGEELFTGVPLFLVELDGVNCHKFSVSGEGDATTYGLTKLFTCTTGKLPVPMPT 60
XX 1 MWSKGEELFTGVPLFLVELDGVNCHKFSVSGEGDATTYGLTKLFTCTTGKLPVPMPT 60
XX
XX 61 LVTLTVYVQCSFRYPDMKQHPFKSAMPEGYVQERTIFFPKDGNVYKTRAEVKEGDTL 120
XX 61 LVTLTVYVQCSFRYPDMKQHPFKSAMPEGYVQERTIFFPKDGNVYKTRAEVKEGDTL 120
XX 61 LVTLTVYVQCSFRYPDMKQHPFKSAMPEGYVQERTIFFPKDGNVYKTRAEVKEGDTL 120
XX
XX 121 VNRHLKIDPFREDENIIGHKLEVYVNSHNYYIMDKQNGIKVNFKIRHNIEDSSVOLA 180
XX 121 VNRHLKIDPFREDENIIGHKLEVYVNSHNYYIMDKQNGIKVNFKIRHNIEDSSVOLA 180
XX 121 VNRHLKIDPFREDENIIGHKLEVYVNSHNYYIMDKQNGIKVNFKIRHNIEDSSVOLA 180
XX
XX Db

```

```

Oy 181 DRYQANTPIGDPVLTLPDNNHYLSTQSALSKDPNKKRPHMLTERTVTAAGITLGMDELKXK 240
Oy 181 DRYQANTPIGDPVLTLPDNNHYLSTQSALSKDPNKKRPHMLTERTVTAAGITLGMDELKXK 240
Oy 241 LSHGFPEVEERQDDGTLPMSCAOSGMDRHPACASARINV 281
Oy 241 LSHGFPEVEERQDDGTLPMSCAOSGMDRHPACASARINV 281
Oy 241 LSHGFPEVEERQDDGTLPMSCAOSGMDRHPACASARINV 281
Oy
Oy RESULT 4
Oy AAB22940
Oy ID AAB22940 standard; protein; 805 AA.
Oy
Oy AC AAB22940;
Oy
Oy DT 10-JAN-2001 (first entry)
Oy
Oy DE GFP-NF-kappa-B fusion protein, SEQ ID NO:178.
Oy
Oy KM Biotector protein; fusion protein; recognition site;
Oy KM cellular targeting sequence; cellular localisation; fluorescent protein;
Oy KM protease activity detection; toxin detection; cellular stress detection;
Oy KM drug discovery; cell based screening.
Oy
Oy XX Aegorrea victoria.
Oy OS Unidentified.
Oy OS Chimeric.
Oy
Oy XX WO200050872-A2.
Oy
Oy PD 31-AUG-2000.
Oy
Oy PF 25-FEB-2000; 2000WO-US004794.
Oy
Oy XX 26-FEB-1999; 99US-0122152P.
Oy PR 08-MAR-1999; 99US-0123399P.
Oy PR 12-JUL-1999; 99US-00352171.
Oy
Oy XX (CELi-) CELLOMICS INC.
Oy
Oy PI Giuliano KA, Kapur R;
Oy
Oy DR MPI; 2000-594086/56.
Oy N-PSDB; AAA93445.
Oy
Oy XX Automated cell-based characterization of toxin by contacting cells
Oy PT containing luminescent reporter molecules with test substance and
Oy PT analyzing optically.
Oy
Oy PS Example 11; Page 329-332; 336pp; English.
Oy
Oy The invention relates to systems, methods and reagents for cell-based
Oy screening or detection of compounds which affect particular biological
Oy functions. The methods of the invention utilise fluorescent biodeceptor
Oy molecules which, when acted on by a compound of interest, cause an
Oy alteration in the cellular distribution of at least the fluorescent
Oy moiety. In one embodiment, the biosensors comprise heat shock proteins
Oy (HSPs) fused to a fluorescent protein (e.g., jellyfish green fluorescent
Oy protein (GFP), or derivatives thereof). Such biosensors are located in
Oy the cytoplasm, but on stress activation translocate to the nucleus. In
Oy another embodiment biodeceptor proteins can be used to detect protease
Oy activity. Such protease biodeceptor fusion proteins comprise one or more
Oy fluorescent proteins; a recognition signal which is cleaved by the
Oy protease; and at least one cellular localisation signal. The latter two
Oy components may be components of a single protein which is acted upon by
Oy the protease, or may be from heterologous sources. Due to the
Oy localisation signal, the biodeceptor protein is localised to a particular
Oy region of the cell. Once acted on by the protease of interest, the
Oy fluorescent protein is cleaved from the localisation sequence, and is
Oy free to migrate to other locations within the cell. The presence of a
Oy second localisation signal attached to the fluorescent protein enables
Oy the fluorescent protein to be directed to a different cellular

```


XX OS Synthetic.
 OS Aequorea victoria.
 OS Homo sapiens.
 XX WO200248693-A1.
 XX PD 20-JUN-2002.
 XX PF 14-DEC-2001; 2001WO-0P010998.
 XX PR 14-DEC-2000; 2000JP-00380327.
 XX PR 30-JAN-2001; 2001JP-00022105.
 XX PA (OLYU) OLYMPUS OPTICAL CO LTD.
 XX PI Kato N, Sakamoto H;
 XX DR WPI; 2002-528176/56.
 XX DR N-PSDB; AAN86125.
 XX PS Autocorrection function-embedded confocal optics-based fluorometric
 PT analyzers for studying behaviors of fluorescence-labeled molecules
 PT particularly intracellular biological molecules like proteins at
 PT molecular level.
 XX Example 2; Page 98-105; 1093p; Japanese.
 CC The invention relates to a fluorometric analyzer that comprises a laser
 CC light source, an optical system for focusing the light beam from laser
 CC light source to a sample to form a confocal region, another optical
 CC system for focusing fluorescence from the sample, a light detector for
 CC measuring intensity, and a recording means to record the variation of
 CC intensity in the measured fluorescence. The analyzers are for studying
 CC behaviors of fluorescence-labeled molecules particularly intracellular
 CC biological molecules like proteins at molecular level, e.g. protein
 CC functions and interactions. Such analyzers are stable and convenient to
 CC operate, thereby enabling easy performance of fluorescence correlation
 CC spectroscopy, fluorescence intensity distribution analysis and
 CC fluorescence represents the amino acid sequence of a fusion protein
 CC comprising a GFP (green fluorescent protein) and hER (human estrogen
 CC receptor) beta protein
 CC XX
 SQ Sequence 793 AA;
 Query Match 85.4%; Score 1280.5; DB 5; Length 793;
 Best Local Similarity 88.4%; Pred. No. 1.4e-122;
 Matches 245; Conservative 2; Mismatches 5; Indels 25; Gaps 1;
 QY 1 MWSKGEELFTGVVPIIVELDGDVNGHKFSVSGEGBDATYGLTLFICTTGKLPVMPPT 60
 Db 1 MWSKGEELFTGVVPIIVELDGDVNGHKFSVSGEGBDATYGLTLFICTTGKLPVMPPT 60
 QY 61 LVTTLTGVVQCFSRYPDHMKQHPFKSAMPEGYVOERTIFFKDDGNYKTRAEVKFGD 120
 Db 61 LVTTLTGVVQCFSRYPDHMKQHPFKSAMPEGYVOERTIFFKDDGNYKTRAEVKFGD 120
 QY 121 VNRTELKGIIDFKEDGNILGHKLEYNYSNHNVIIMADKQKGIKVNFKIRHNIEDGSVOLA 180
 Db 121 VNRTELKGIIDFKEDGNILGHKLEYNYSNHNVIIMADKQKGIKVNFKIRHNIEDGSVOLA 180
 QY 181 DHYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHNVLEFVTAAGITLGMDELYKK 240
 Db 181 DHYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHNVLEFVTAAGITLGMDELYKK 240
 QY 241 LSHRPFPEVEBEQDDGTLPMSCAQSSGMDRHPAAQASA 277
 Db 240 -----SGLRSRAQACASS 252

ID AAN85037 standard; protein; 1171 AA.
 XX AAN85037;
 AC 17-OCT-2003 (revised)
 DT 08-FEB-1999 (first entry)
 XX Green fluorescent protein-NFAT1 fusion product.
 DE Human, NFAT1 gene; fusion protein; green fluorescent protein; GFP,
 KW intracellular signalling; chimera.
 KW Aequorea victoria.
 OS Homo sapiens.
 OS Chimeric.
 XX WO9845704-A2.
 XX PD 15-OCT-1998.
 XX PF 07-APR-1998; 98WO-DK00145.
 XX PR 07-APR-1997; 97DK-00000392.
 XX PA (NOVO) NOVO-NORDISK AS.
 XX PI Thastrup O, Petersen Bjorn S, Tullin S, Kasper A, Scudder K;
 XX DR WPI; 1998-594491/50.
 XX DR N-PSDB; AAV11037.
 XX PT Determining effect on signalling pathways in live cells from
 PT redistribution of luminoephores - specifically fusions of green
 PT fluorescent protein with a signalling component, and new apparatus,
 PT particularly for identifying toxins and potential therapeutic agents.
 XX Example 19; Page 258-260; 326p; English.
 XX The present sequence represents a green fluorescent protein (GFP)-human
 CC NFAT1 fusion protein. The fusion protein is used in an assay to exemplify
 CC the invention. The specification describes how quantitative information
 CC about the influence of a molecule on a cellular response is obtained by
 CC recording the variation, caused by the molecule, on mechanically intact
 CC living cells, in the spatially distributed light emitted from a
 CC luminoephore present in the cells. The variation in light emission is
 CC processed to provide information that correlates spatial distribution to
 CC the degree of the molecule. The method is used to identify agents that
 CC (in)directly affect intracellular signalling, especially to screen for
 CC potential therapeutic agents or toxins, and to identify new drug targets.
 CC (Updated on 17-OCT-2003 to standardise OS field)
 CC XX
 SQ Sequence 1171 AA;
 Query Match 85.4%; Score 1280.5; DB 2; Length 1171;
 Best Local Similarity 93.9%; Pred. No. 2.5e-122;
 Matches 245; Conservative 1; Mismatches 6; Indels 9; Gaps 2;
 QY 1 MWSKGEELFTGVVPIIVELDGDVNGHKFSVSGEGBDATYGLTLFICTTGKLPVMPPT 60
 Db 1 MWSKGEELFTGVVPIIVELDGDVNGHKFSVSGEGBDATYGLTLFICTTGKLPVMPPT 60
 QY 61 LVTTLTGVVQCFSRYPDHMKQHPFKSAMPEGYVOERTIFFKDDGNYKTRAEVKFGD 120
 Db 61 LVTTLTGVVQCFSRYPDHMKQHPFKSAMPEGYVOERTIFFKDDGNYKTRAEVKFGD 120
 QY 121 VNRTELKGIIDFKEDGNILGHKLEYNYSNHNVIIMADKQKGIKVNFKIRHNIEDGSVOLA 180
 Db 121 VNRTELKGIIDFKEDGNILGHKLEYNYSNHNVIIMADKQKGIKVNFKIRHNIEDGSVOLA 180
 QY 181 DHYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHNVLEFVTAAGITLGMDELYKK 240
 Db 181 DHYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHNVLEFVTAAGITLGMDELYKK 240

QY 241 LSHGF-----PPEVEEDDG 255
 DB 241 ---GLRSRAMNABERQPDG 258

RESULT 8
 ABP72573
 ID ABP72573 standard; protein; 357 AA.
 XX ABP72573;
 XX
 DT 23-OCT-2003 (revised)
 DT 29-MAY-2003 (first entry)
 XX
 DE GFP-Drosophila Melted protein pleckstrin homology domain fusion.
 XX
 KM Melted; dmelted; pleckstrin; phosphatidylinositol 5-phosphate; PI(5)P;
 XX binding protein; green fluorescent protein; GFP.
 XX
 OS Aequorea victoria.
 OS Drosophila melanogaster.
 OS Chimeric.

Key Location/Qualifiers
 FT Region 1..245
 FT /label= GFP
 FT 253..350
 FT /label= PH
 FT /note= "dmelted pleckstrin homology domain"

MO2003014152-A2.
 PD 20-FEB-2003.
 XX
 PF 08-AUG-2002; 2002WO-1B003645.
 XX
 PR 09-AUG-2001; 2001GB-00019446.
 XX
 PA (EURO-) EURO MOLECULAR BIOLOGY LAB.
 XX
 PI Cohen S, Teleman A;
 XX
 DR WPI; 2003-256540/25.
 XX
 PT New proteins that bind phosphatidylinositol 5-phosphate (PI(5)P), useful
 PT for treating a disease or disorder in which abnormal levels of PI(5)P are
 PT implicated, or visualizing or measuring PI(5)P in vivo or in vitro.
 XX
 PS Example; Page 22; 41pp; English.

CC The present sequence is the protein sequence of a fusion protein between
 CC green fluorescent protein (GFP) and the pleckstrin homology (PH) domain
 CC (see also ABP72573) of the Melted protein (dmelted) of Drosophila
 CC melanogaster. Also included in the sequence are short peptide linkers at
 CC the N- and C-termini of the PH moiety. The fusion protein was constructed
 CC and expressed in Drosophila S2 cells in an example from the invention.
 CC The invention relates to proteins that bind to phosphatidylinositol 5-
 CC phosphate (PI(5)P). The PI(5)P binding protein preferably comprises a PH
 CC domain of a Melted protein derived from a vertebrate, a metazoan or an
 CC insect. The binding protein, or a fusion protein comprising it, is useful
 CC for treating a disease or disorder in which abnormal levels of PI(5)P are
 CC implicated, such as diseases that result from dysfunctional signal
 CC transduction and which affect gene expression and cellular proliferation
 CC and differentiation. The PI(5)P binding protein is also useful for
 CC purifying a PI(5)P-containing membrane fraction, or visualizing the
 CC subcellular location of PI(5)P in a cell. Fluorescence microscopy of
 CC transfected Drosophila cells revealed a cortical localisation of the GFP-
 CC dmelted PH fusion construct, showing that PI(5)P is located cortically in
 CC cells. (Updated on 23-Oct-2003 to standardise OS field)

Sequence 357 AA;
 SQ
 Query Match 85.3%; Score 1279.5; DB 6; Length 357;

Best Local Similarity 94.6%; Pred. No. 5.4e-123;
 Matches 245; Conservative 1; Mismatches 6; Indels 7; Gaps 2;

QY 1 MWSKGEELFTGVPLIVELDGVNGHKPSVSGEGDATTGKLTIKFICTTGKLPVMPPT 60
 DB 1 MWSKGEELFTGVPLIVELDGVNGHKPSVSGEGDATTGKLTIKFICTTGKLPVMPPT 60

QY 61 LVTLITVGVCFSRYPDHMKQDFFKSNAPBGYYOERTIFFKDDGNNYTRAEVFEEDTL 120
 DB 61 LVTLITVGVCFSRYPDHMKQDFFKSNAPBGYYOERTIFFKDDGNNYTRAEVFEEDTL 120

QY 121 VNRILKGIIDFKEDGNILGHLELYNNSHNYIADKQNGIKVNFIRHNIEDGSVOLA 180
 DB 121 VNRILKGIIDFKEDGNILGHLELYNNSHNYIADKQNGIKVNFIRHNIEDGSVOLA 180

QY 181 DHYQONTPIGDGPVLLPNNHYLSTQSALSKDPNEKRDMLLEFYTAGITLGMDELTKK 240
 DB 181 DHYQONTPIGDGPVLLPNNHYLSTQSALSKDPNEKRDMLLEFYTAGITLGMDELTKK 240

QY 241 --LSHGF-----PPEVEEDG 252
 DB 241 GLRSRGLNGSNQPMIEGQ 259

RESULT 9
 AAW85032
 ID AAW85032 standard; protein; 997 AA.
 XX
 XX AAW85032;
 AC
 DT 17-OCT-2003 (revised)
 DT 08-FEB-1999 (first entry)
 XX
 DE Green fluorescent protein-IkappaB kinase alpha subunit fusion product.
 XX
 KM Human; IkappaB kinase gene; alpha subunit; fusion protein;
 KM green fluorescent protein; GFP; intracellular signalling; chimera.
 XX
 OS Aequorea victoria.
 OS Homo sapiens.
 OS Chimeric.

PN WO9845704-A2.
 XX
 PD 15-OCT-1998.
 PD
 XX 07-APR-1998; 98WO-DK000145.
 PF
 XX 07-APR-1997; 97DK-00000392.
 PR
 XX (NOVO) NOVO-NORDISK AS.
 PA
 PI Thastrup O, Petersen Bjorn S, Tullin S, Kasper A, Scudder K;
 PI N-PSDB; AAV71077.
 DR WPI; 1998-594491/50.
 XX
 PT Determining effect on signalling pathways in live cells from
 PT redistribution of lumino-phores - specifically fusions of green
 PT fluorescent protein with a signalling component, and new apparatus,
 PT particularly for identifying toxins and potential therapeutic agents.

PS Example 11; Page 231-233; 326pp; English.

CC The present sequence represents a green fluorescent protein (GFP)-human
 CC IkappaB kinase alpha subunit fusion protein. The fusion protein is used
 CC in an assay to exemplify the invention. The specification describes how
 CC quantitative information about the influence of a molecule on a cellular
 CC response is obtained by recording the variation, caused by the molecule,
 CC on mechanically intact living cells, in the spatially distributed light
 CC emitted from a lumino-phore present in the cells. The variation in light
 CC emission is processed to provide information that correlates spatial
 CC distribution to the degree of the molecule. The method is used to

CC identify agents that (in)directly affect intracellular signalling, and to
 CC especially to screen for potential therapeutic agents or toxins, and to
 CC identify new drug targets. (Updated on 17-OCT-2003 to standardise OS
 CC field)

CC Sequence 997 AA;

Query Match 85.3%; Score 1279; DB 2; Length 997;
 Best Local Similarity 90.4%; Pred. No. 2,9e-122; Indels 14; Gaps 1;
 Matches 245; Conservative 2; Mismatches 10;

QY 1 MWSKGEELFTGVPLIVELDGVNGHKFSVSGEGGATYGLTLKFTCTTGKLPVMPPT 60
 DB 1 MWSKGEELFTGVPLIVELDGVNGHKFSVSGEGGATYGLTLKFTCTTGKLPVMPPT 60
 QY 61 LVTTITTVGVCFSRYPDHMKOHDFPKSAMPEGYVOERTIFFKDDGNTYKTRAIVKREGDTL 120
 DB 61 LVTTITTVGVCFSRYPDHMKOHDFPKSAMPEGYVOERTIFFKDDGNTYKTRAIVKREGDTL 120
 QY 121 VNRTELKIDPKEDGNILGHKLEYNNSHNYIYIMADKQNGIKVNFKIRHNIEDGSVOLA 180
 DB 121 VNRTELKIDPKEDGNILGHKLEYNNSHNYIYIMADKQNGIKVNFKIRHNIEDGSVOLA 180
 QY 181 DHYOQNTPIGDGPVLLPDNHYLSTOSALSCKDPNEKRDMVLEFVTAAGITLGMDELTK 240
 DB 181 DHYOQNTPIGDGPVLLPDNHYLSTOSALSCKDPNEKRDMVLEFVTAAGITLGMDELTK 240
 QY 241 LSHGFPPEVEBDDGTLPMSCAQESGMDRHP 271
 DB 240 -----SGLSRGAQASNSTMERPP 257

RESULT 10

AAW85010
 ID AAW85010 standard; protein; 607 AA.

AC AAW85010;

DT 17-OCT-2003 (revised)
 DT 08-FEB-1999 (first entry)

DE p38-green fluorescent protein fusion product.

KW Human; p38 gene; fusion protein; green fluorescent protein; GFP;
 KW intracellular signalling; chimera.

OS Aequorea victoria.

OS Homo sapiens.

OS Chimeric.

PM W09845704-A2.

PD 15-OCT-1998.

PF 07-APR-1998; 98WO-DK000145.

PR 07-APR-1997; 97DK-00000392.

PA (NOVO) NOVO-NORDISK AS.

PI Thastrup O, Petersen Bjorn S, Tullin S, Kasper A, Scudder K;

DR WPI; 1998-594491/50.

DR N-PSDB; AAV71025.

PT Determining effect on signalling pathways in live cells from
 PT redistribution of lumiphores - specifically fusions of green
 PT fluorescent protein with a signalling component, and new apparatus,
 PT particularly for identifying toxins and potential therapeutic agents.

PS Example 8; Page 90-91; 326pp; English.

XX The present sequence represents a human p38-green fluorescent protein

CC fusion product. The fusion protein is used in an assay that exemplifies
 CC the invention. The specification describes how quantitative information
 CC about the influence of a molecule on a cellular response is obtained by
 CC recording the variation, caused by the molecule, on mechanically intact
 CC living cells, in the spatially distributed light emitted from a
 CC lumiphore present in the cells. The variation in light emission is
 CC processed to provide information that correlates spatial distribution to
 CC the degree of the molecule. The method is used to identify agents that
 CC (in)directly affect intracellular signalling, especially to screen for
 CC potential therapeutic agents or toxins, and to identify new drug targets.
 CC (Updated on 17-OCT-2003 to standardise OS field)

CC Sequence 607 AA;

Query Match 85.2%; Score 1278; DB 2; Length 607;
 Best Local Similarity 84.7%; Pred. No. 1.7e-122; Indels 24; Gaps 3;
 Matches 249; Conservative 6; Mismatches 15;

QY 1 MWSKGEELFTGVPLIVELDGVNGHKFSVSGEGGATYGLTLKFTCTTGKLPVMPPT 60
 DB 1 MWSKGEELFTGVPLIVELDGVNGHKFSVSGEGGATYGLTLKFTCTTGKLPVMPPT 60
 QY 61 LVTTITTVGVCFSRYPDHMKOHDFPKSAMPEGYVOERTIFFKDDGNTYKTRAIVKREGDTL 120
 DB 61 LVTTITTVGVCFSRYPDHMKOHDFPKSAMPEGYVOERTIFFKDDGNTYKTRAIVKREGDTL 120
 QY 121 VNRTELKIDPKEDGNILGHKLEYNNSHNYIYIMADKQNGIKVNFKIRHNIEDGSVOLA 180
 DB 121 VNRTELKIDPKEDGNILGHKLEYNNSHNYIYIMADKQNGIKVNFKIRHNIEDGSVOLA 180
 QY 181 DHYOQNTPIGDGPVLLPDNHYLSTOSALSCKDPNEKRDMVLEFVTAAGITLGMDELTK 239
 DB 181 DHYOQNTPIGDGPVLLPDNHYLSTOSALSCKDPNEKRDMVLEFVTAAGITLGMDELTK 240
 QY 240 -----KLSHGFP-----EVEBDDGTLPMSCAQESGMDRHPAACA 276
 DB 241 GLRSRGMKSCRPFTFYRQELNKTITWEVERYONLSPV-----SGAGAYVCAA 288

RESULT 11

AAW85574
 ID AAW85574 standard; protein; 2608 AA.

AC AAW85574;

DT 07-JUN-2000 (first entry)

DE Hs-UNC-53/3/GFP fusion insert of plasmid pG13303.

KW UNC-53; Caenorhabditis elegans; microtubule; neural regeneration;
 KW anticancer; anti-neurodegeneration; antifibrotic; anti-adhesive; human;
 KW antisclerotic; antimetabolic; anti-arthritis; autoimmune disease.

OS Synthetic.

OS Homo sapiens.

OS Key

PM Misc-difference 1194
 PT Location/Qualifiers
 PT /note= "unspecified"

PI W09963080-A1.

PD 09-DEC-1999.

PF 02-JUN-1999; 99WO-EP003848.

PR 03-JUN-1998; 98GB-00011962.

PA (JANCO) JANSSEN PHARM NV.

PI Luyten WHM, De Raeymaeker MC, Geysen JGH, Bogaert THOE;
 PI Maerten LJS, Verhaesele P, Van De Craen M;

XX

DR WPI; 2000-116370/10.
 DR N-PSDB; AAA07846.
 XX
 XX Novel proteins and nucleic acids e.g. for treating neurodegeneration.
 PS Disclousure; Fig 7e; 146pp; English.
 XX
 XX The invention provides vertebrate (human) protein homologue of a UNC-53
 CC protein of *Caenorhabditis elegans*. The UNC-53 binds to microtubules or
 CC their plus ends. The UNC-53 sequences are used to promote neural
 CC regeneration, revascularization and wound healing; also for treating
 CC neurodegenerative disease, acute traumatic injury, fibrotic disease and
 CC autoimmune diseases (e.g. rheumatoid arthritis and sclerosis). The UNC-53
 CC polynucleotides can be used for recombinant production of the proteins,
 CC as a source of probes for detecting allelic variants and polymorphisms,
 CC for sequencing genomic DNA and for detecting UNC-53 expression; and as
 CC source of therapeutic antisense sequences. Cells that express the protein
 CC are used to identify regulators of cell shape, growth, motility and
 CC migration. They can also be used to identify proteins that are involved
 CC in signal transduction pathways also involving UNC-53, and to identify
 CC compounds that alter attachment of UNC-53 to microtubules. A target gene
 CC coupled to a UNC-53 encoding sequence may be used to deliver the target
 CC gene to a cellular microtubule or its plus ends. The present sequence
 CC represents the amino acid sequence of a full-length Hs-UNC-53/3 in fusion
 CC with GFP insert of plasmid pG13305
 CC
 XX
 SQ Sequence 2608 AA;
 Query Match 85.2%; Score 1278; DB 3; Length 2608;
 Best Local Similarity 91.4%; Pred. No. 1,5e-121; Mismatches 5; Gaps 2;
 Matches 245; Conservative 5; Indels 6; Gaps 2;
 QY 1 MVSKEELFTGVVPIIVELDGVNKGKFSVSGEGDATTGKLTLLKFICTTGKLPVPMPT 60
 Db 1 MVSKEELFTGVVPIIVELDGVNKGKFSVSGEGDATTGKLTLLKFICTTGKLPVPMPT 60
 QY 61 LVTTLTLYGQCFSRYPDMKQHDFFKSAMPEGYVOERTIFFKDDGNYKTRAVKFGDTL 120
 Db 61 LVTTLTLYGQCFSRYPDMKQHDFFKSAMPEGYVOERTIFFKDDGNYKTRAVKFGDTL 120
 QY 121 VNRLEKGIKDFKEDGNILGHKLEYNNSHNYITMADKQNGIKVNFIRNIEDGSVOLA 180
 Db 121 VNRLEKGIKDFKEDGNILGHKLEYNNSHNYITMADKQNGIKVNFIRNIEDGSVOLA 180
 QY 181 DHYQONTPIGDGPVLLPDNHYLSTQSALSKDNEKRDHVVLEFYTAAGITLGMDELYK 239
 Db 181 DHYQONTPIGDGPVLLPDNHYLSTQSALSKDNEKRDHVVLEFYTAAGITLGMDELYK 240
 QY 240 -KLSH---GFPPEVEQDDGTPMCA 262
 Db 241 SDLEHMPVLGVASKRQPAVGSKRVHTA 268
 Db 241 SDLEHMPVLGVASKRQPAVGSKRVHTA 268
 RESULT 12
 ADB90657
 ID ADB90657 standard; protein; 1089 AA.
 XX
 XX ADB90657;
 AC
 XX
 XX 04-DEC-2003 (first entry)
 DT
 XX
 XX Human and rat chimeric protein #2.
 DE
 XX
 XX nuclear translocation protein; glucocorticoid receptor DNA; DNA chimera;
 KW Cytostatic; Cardiant; Antiarteriosclerotic; Antiarthritis;
 KM Antiinflammatory; Gene therapy; cancer; heart disease; arteriosclerosis;
 KM arthritis; inflammatory states; human; rat.
 XX
 XX Chimeric.
 OS Homo sapiens.
 OS Rattus sp.
 XX
 XX
 PN MO2003027639-A2.

XX
 PD 03-APR-2003.
 XX
 XX 30-SEP-2002; 2002MO-US031043.
 PF
 XX 28-SEP-2001; 2001US-0325178P.
 PR
 XX (HAGER/) HAGER G L.
 PA
 XX Hager GL, Mackem S;
 PI
 XX WPI; 2003-402984/38.
 DR
 XX Making a recombinant nuclear translocation protein, useful for treating
 PT cancer, heart disease, by covalently connecting glucocorticoid receptor
 PT DNA, superfamily receptor DNA and nucleic acid sequence for a marker
 PT protein domain.
 PS Disclousure; Page 75-81; 115pp; English.
 XX
 XX The present sequence relates to making a recombinant nuclear
 CC translocation protein by covalently connecting a glucocorticoid receptor
 CC DNA sequence coding for the cytoplasmic/nuclear translocation domain of
 CC the glucocorticoid receptor protein, a superfamily receptor DNA sequence
 CC coding for the ligand binding domain of the superfamily receptor protein,
 CC and a nucleic acid sequence for a marker protein domain, forming a DNA
 CC chimera. The methods, chimera, and protein are useful for treating
 CC defective translocation of a superfamily receptor protein from the
 CC cytoplasm to the nucleus of the cell in an animal, such as cancer, heart
 CC disease, arteriosclerosis, arthritis, and inflammatory states. The
 CC present sequence represents a rat/human chimera of the invention.
 CC
 XX
 SQ Sequence 1089 AA;
 Query Match 85.1%; Score 1277; DB 7; Length 1089;
 Best Local Similarity 86.4%; Pred. No. 5.2e-122; Mismatches 25; Indels 6; Gaps 2;
 Matches 247; Conservative 8; Gaps 2;
 QY 1 MVSKEELFTGVVPIIVELDGVNKGKFSVSGEGDATTGKLTLLKFICTTGKLPVPMPT 60
 Db 1 MVSKEELFTGVVPIIVELDGVNKGKFSVSGEGDATTGKLTLLKFICTTGKLPVPMPT 60
 QY 61 LVTTLTLYGQCFSRYPDMKQHDFFKSAMPEGYVOERTIFFKDDGNYKTRAVKFGDTL 120
 Db 61 LVTTLTLYGQCFSRYPDMKQHDFFKSAMPEGYVOERTIFFKDDGNYKTRAVKFGDTL 120
 QY 121 VNRLEKGIKDFKEDGNILGHKLEYNNSHNYITMADKQNGIKVNFIRNIEDGSVOLA 180
 Db 121 VNRLEKGIKDFKEDGNILGHKLEYNNSHNYITMADKQNGIKVNFIRNIEDGSVOLA 180
 QY 181 DHYQONTPIGDGPVLLPDNHYLSTQSALSKDNEKRDHVVLEFYTAAGITLGMDELYK 240
 Db 181 DHYQONTPIGDGPVLLPDNHYLSTQSALSKDNEKRDHVVLEFYTAAGITLGMDELYK 240
 QY 241 -----LSHGFPPEVEQDDGTPMCAQESGMDRHPACASRIIV 281
 Db 241 GALDQSKESLAPPGRDEVGSL-LGQGGSVYVDFYKSLRGATVAV 285
 Db 241 GALDQSKESLAPPGRDEVGSL-LGQGGSVYVDFYKSLRGATVAV 285
 RESULT 13
 ADE24110
 ID ADE24110 standard; protein; 284 AA.
 XX
 XX ADE24110;
 AC
 XX
 XX 29-JAN-2004 (first entry)
 DT
 XX
 XX Discosoma sp DsRed1 reporter protein #2.
 DE
 XX
 XX site-specific DNA recombination; Cre recombinase;
 KW protein-transduction domain; gene activation; gene inactivation;
 KW chromosomal translocation; reporter protein.
 XX
 XX

OS Synthetic.
OS Discosoma sp.
XX WO2003070931-A2.
XX 28-AUG-2003.
XX 19-FEB-2003; 2003WO-EP001680.
XX 21-FEB-2002; 2002DE-01007313.
XX 16-JUL-2002; 2002DE-01032196.
XX (VISI-) VISION 7 GMBH.
XX
XX Baum C, Will E, Osterlag W, Klump H, Schiedlmeier B;
XX WPI; 2003-767353/72.
XX
XX In vitro or in vivo site-specific DNA recombination, useful e.g. for gene
XX inactivation, using Cre recombinase that lacks heterologous protein-
XX transduction domain.
XX
XX Claim 22; SEQ ID NO 7; 84bp; German.
XX
XX This invention describes a novel method of site-specific DNA
XX recombination in eukaryotic cells in vitro comprising using a
XX bacteriophage P1 Cre recombinase that lacks heterologous protein-
XX transduction domains. The invention also contains a reporter system for
XX detecting site-specific DNA recombination in eukaryotic cells.
XX Recombination with Cre recombinase lacking heterologous protein-
XX transduction domains is used for therapeutic site-specific recombination
XX in eukaryotic cells, in vivo or in vitro, e.g. for partial inactivation
XX or activation of genes, also to create a chromosomal translocation. The
XX modified recombinase can enter intact (not electrically or chemically
XX treated) cells when added to culture medium and catalyzes recombination
XX events in the nucleus with recombination rate over 50%, after a single
XX application. The Cre-recombinase does not need to be supplied from
XX nucleic acid introduced into the cell, so its effect is immediate but
XX short-lived, limiting cytotoxic effects; eliminating the risk of non-
XX specific integration of DNA into the genome and providing quick results.
XX This sequence represents a Discosoma sp reporter protein DsRed1 which
XX contains a fragment of green fluorescent protein (eGFP) used in the
XX method described in the invention.
XX
XX Sequence 284 AA;
SQ
Query Match 85.1%; Score 1276.5; DB 7; Length 284;
Best Local Similarity 94.2%; Pred. No. 7.8e-123;
Matches 243; Conservative 2; Mismatches 10; Indels 3; Gaps 1;
QY 2 VSKGEELFTGVVPIIVELDGVNGHKFSVSGEGGDATYKLTIKFICTTGKLPVMPPTL 61
DB 18 VSKGEELFTGVVPIIVELDGVNGHKFSVSGEGGDATYKLTIKFICTTGKLPVMPPTL 77
QY 62 VTTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNKXTRAAYVFBGDTLV 121
DB 78 VTTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNKXTRAAYVFBGDTLV 137
QY 122 NRILKGIIDFKEDGNILGHKLEYNNSHNVIYIMADKQNGIKVNFKIRHNIEDSVQOLAD 181
DB 138 NRILKGIIDFKEDGNILGHKLEYNNSHNVIYIMADKQNGIKVNFKIRHNIEDSVQOLAD 197
QY 182 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDMHVLLEFVTAAGITLGMDELXXK - 239
DB 198 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDMHVLLEFVTAAGITLGMDELXXK 257
QY 240 -KLSHGFPPEVEBDDGT 256
DB 258 LRSSADNPASTTNKDAIG 275
RESULT 14
ADB90661

ID ADB90661 standard; protein; 1090 AA.
XX
XX AC ADB90661;
XX
XX 04-DEC-2003 (first entry)
XX
XX Human and rat chimeric protein #4.
XX
XX nuclear translocation protein; glucocorticoid receptor DNA; DNA chimera;
XX Cycloastic; Cardiant; Antiarteriosclerotic; Antiarthritic;
XX Antiinflammatory; Gene therapy; cancer; heart disease; arteriosclerosis;
XX arthritis; inflammatory states; human; rat.
XX
XX Chimeric.
XX OS Homo sapiens.
XX OS Rattus sp.
XX
XX WO2003027639-A2.
XX
XX 03-APR-2003.
XX
XX 30-SEP-2002; 2002WO-US031043.
XX
XX 28-SEP-2001; 2001US-0325178P.
XX
XX (HAGE/) HAGER G L.
XX
XX Hager GL, Mackem S;
XX WPI; 2003-402984/38.
XX N-PSDB; ADB90660.
XX
XX Making a recombinant nuclear translocation protein, useful for treating
XX cancer, heart disease, by covalently connecting glucocorticoid receptor
XX DNA, superfamily receptor DNA and nucleic acid sequence for a marker
XX protein domain.
XX
XX Disclosure; Page 108-114; 115bp; English.
XX
XX The present sequence relates to making a recombinant nuclear
XX translocation protein by covalently connecting a glucocorticoid receptor
XX DNA sequence coding for the cytoplasmic/nuclear translocation domain of
XX the glucocorticoid receptor protein, a superfamily receptor DNA sequence
XX coding for the ligand binding domain of the superfamily receptor protein,
XX and a nucleic acid sequence for a marker protein domain, forming a DNA
XX chimera. The methods, chimera, and protein are useful for treating
XX defective translocation of a superfamily receptor protein from the
XX cytoplasm to the nucleus of the cell in an animal, such as cancer, heart
XX disease, arteriosclerosis, arthritis, and inflammatory states. The
XX present sequence represents a rat/human chimera of the invention.
XX
XX Sequence 1090 AA;
SQ
Query Match 85.1%; Score 1276.5; DB 7; Length 1090;
Best Local Similarity 86.1%; Pred. No. 5.9e-122;
Matches 247; Conservative 8; Mismatches 25; Indels 7; Gaps 2;
QY 1 MVSKEELFTGVVPIIVELDGVNGHKFSVSGEGGDATYKLTIKFICTTGKLPVMPPT 60
DB 1 MVSKEELFTGVVPIIVELDGVNGHKFSVSGEGGDATYKLTIKFICTTGKLPVMPPT 60
QY 61 LVTTLYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNKXTRAAYVFBGDTLV 120
DB 61 LVTTLYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNKXTRAAYVFBGDTLV 120
QY 121 VNRILKGIIDFKEDGNILGHKLEYNNSHNVIYIMADKQNGIKVNFKIRHNIEDSVQOLA 180
DB 121 VNRILKGIIDFKEDGNILGHKLEYNNSHNVIYIMADKQNGIKVNFKIRHNIEDSVQOLA 180
QY 181 DHYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDMHVLLEFVTAAGITLGMDELXXK 240
DB 181 DHYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDMHVLLEFVTAAGITLGMDELXXK 240

```

QY      241 -----LSHGFPPEVEEQDDGTLPMSCAQESGMDRHPAACASARINV 281
           |      :: | : | : | : | : | : | : | : | : | : | : |
Db      241 GAILDSKESLAPGRDEVPGSL-LGQGRGSVMDPFYSLRGGATVKV 286

```

RESULT 1.5

ID	ADE28562	standard; protein; 286 AA.
XY		

AC ADE28562;

DT 29-JAN-2004 (first entry)
 VV

EGFP/ human CD1a fusion construct protein.

KM cytotoxic; antidiabetic; antiinflammatory; dermatological;
 KM CD1 fusion protein; endosomal targeting; antibacterial; vituocide;
 KM viral infection; autoimmune disorder; diabetes; lupus;
 KM immunosuppressive; neuroprotective; immune response; microbial;
 KM multiple sclerosis; cancer; gene therapy; vaccine; human;
 KM chimeric; enhanced green fluorescent protein.

Unidentified.

OS Homo sapiens.

US Homo sapiens.
VY

PN WO2003066820-A2.

PD 14-AUG-2003.

05-FEB-2003; 2003WO-US003550.

05-FEB-2002; 2002US-0355432P.

PA (REGC) UNIV CALIFORNIA

Modlin RL;

DR WPI; 2003-748121/70.

DK N-PSDB; ADE28561.

PT new chimeric nucleic acid molecule encoding a CD1 fusion protein, useful
PT for inducing an immune response against antigens associated with, e.g.
PT microbial or viral infections, autoimmune disorders or cancer.

Claim 26; Fig 11A; 114pp; English.

The invention relates to a novel chimeric nucleic acid molecule consisting of a nucleotide sequence encoding a CD1 fusion protein comprising a CD1 endosomal targeting sequence or its fragment and an antigen of interest. The polynucleotide of the invention demonstrates antibacterial, virucidal, cytostatic, antidiabetic, antiinflammatory, dermatological, immunosuppressive and neuroprotective activities whilst the composition and methods may be useful for inducing an immune response against antigens of interest which are associated with microbial or viral infections, autoimmune disorders including diabetes, lupus and multiple sclerosis and antigens associated with cancers. Furthermore, the composition may be utilised in gene therapy and vaccine development. The current sequence is that of the BSPF (enhanced green fluorescent protein) / human CD1a fusion construct protein of the invention.

Sequence 286 AA;
SQ

Query Match	85.1%	Score 1276;	DB 7;	Length 286;
Best Local Similarity	99.2%;	Pred. No. 8.8e-123;		
Matches 239; Conservative	1;	Mismatches	1;	Indels 0; Gaps 0

QY 1 MWSKSEBELFTGVPLIVELVDGVDNHHKRSVSGSEGGDATTGTLTKPLCTTGKLPVMPMT 60

Db 20 MWSKSEBELFTGVPLIVELVDGVDNHHKRSVSGSEGGDATTGTLTKPLCTTGKLPVMPMT 79

QY 61 LVYTLTYGVQCSRYPDHMKODFFKSMPEGVQERTIFPKDDGNVYTRAEVFEEDTL 120

Db	80	LVTTLVGYQCSRYPDHMKQHDFFKSAMPGYQBERTIFFKODGNKKTAEVXFEGDTL	139
QY	121	VNRLEKGIDEFEDGNILGHKLEYNVNSHNYIMADKQNGIKVNFKIRNIEDGSVOLA	180
Db	140	VNRLEKGIDEFEDGNILGHKLEYNVNSHNYIMADKQNGIKVNFKIRNIEDGSVOLA	199
QY	181	DHYQONTPIGDGPVLLPNHNLSTOSALSCKDPNKRDMVLLEVTYAGITLGMDELYKK	240
Db	200	DHYQONTPIGDGPVLLPNHNLSTOSALSCKDPNKRDMVLLEVTYAGITLGMDELYKS	259
QY	241	L 241	
Db	260	I 260	

Search completed: May 14, 2004, 08:25:16
Job time : 63 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 14, 2004, 08:23:58 / Search time 21 Seconds
(without alignments)
1287,134 Million cell updates/sec

Title: US-09-931-232-1

Perfect score: 1500
Sequence: 1 MVSKEBELFTGVPIIVLELD.....AQESGMDRHPACASARINW 281

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR78:*
2: PIR1:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1245	83.0	238	1 JQ1514	green-fluorescent
2	221.5	14.8	461	1 DCM50	ornithine decarbox
3	221.5	14.8	461	2 I56477	ornithine decarbox
4	210.5	14.0	461	2 I55356	ornithine decarbox
5	208.5	13.9	461	1 DCRTO	ornithine decarbox
6	200	13.3	455	1 DCHYOC	ornithine decarbox
7	154.5	10.3	461	1 DCHHO	ornithine decarbox
8	111.5	7.4	450	1 DCHHO	ornithine decarbox
9	103	6.9	460	2 A43563	ornithine decarbox
10	101	6.7	785	2 H72228	hypothetical prote
11	99.5	6.6	1224	1 ERUDAH	coaromer complex a
12	95.5	6.4	4307	2 KXRTCI	hypothetical prote
13	92.5	6.2	752	1 KXRTCI	protoprotein convert
14	92	6.1	336	2 C64468	hypothetical prote
15	91	6.1	357	2 G81355	RNA (uracil-5-)-m
16	91	6.0	632	2 T06586	DNA-binding protei
17	90.5	6.0	861	2 H64102	leucine-tRNA 19as
18	90.5	6.0	1193	2 F68698	DNA-directed RNA p
19	90	6.0	1039	2 H95115	conserved hypobes
20	89	5.9	433	2 B90103	synaptogamin o-p65
21	88.5	5.9	502	2 JH0414	T-complex protein
22	88.5	5.9	578	1 I40794	diacylglycerol
23	88.5	5.9	887	2 E82590	leucyl-tRNA synthet
24	88.5	5.9	889	2 JCS576	inter-alpha-trypsi
25	88.5	5.9	6642	2 T27577	protein UNC-89 - C
26	88	5.9	471	2 T27856	hypothetical prote
27	88	5.9	1259	2 A81055	probable exported
28	87.5	5.8	2573	2 D71614	hypothetical prote
29	87	5.8	281	2 AD2052	hypothetical prote

30	86.5	5.8	370	2 E70390	iron-sulfur cofact
31	86.5	5.8	461	2 T06936	photosystem II chl
32	86.5	5.8	860	2 AC0582	leucyl-tRNA synthet
33	86.5	5.8	874	2 JQ4930	S-layer protein pr
34	86.5	5.8	2222	1 A36028	DNA-directed DNA p
35	86	5.7	263	2 S53488	water-stress-induc
36	86	5.7	1039	2 D97985	hypothetical prote
37	86	5.7	1290	2 S76853	hypothetical prote
38	85.5	5.7	625	2 D83917	DNA topoisomerase
39	85.5	5.7	788	1 JQVLDH	DNA-directed DNA p
40	85.5	5.7	877	2 AC2211	home transport pro
41	85.5	5.7	1134	2 A60234	IGA Fc receptor pr
42	85.5	5.7	1164	1 FCS0AG	IGA Fc receptor pr
43	85	5.7	331	2 T41758	AcAMPV orf13 - Bom
44	85	5.7	466	2 B64207	heterocyst maturat
45	85	5.7	3744	2 S46715	hypothetical prote

ALIGNMENTS

RESULT 1

JQ1514

green-fluorescent protein [validated] - hydromedusa (Aequorea victoria)

C/Species: Aequorea victoria

C/Date: 03-Dec-1999 #sequence revision 03-Dec-1999 #ext change 23-Mar-2001

C/Accession: US0692; JQ1514; F00335; S48693; S51330; S51331

R/Prasner, D.C.; Eckenrode, V.K.; Ward, W.W.; Prendergast, F.G.; Cormier, M.J.
Gene 111, 229-233, 1992

A/Title: Primary structure of the Aequorea victoria green-fluorescent protein.

A/Reference number: JQ1514; MUID:92175527; PMID:1347277

A/Accession: J06092

A/Molecule type: DNA

A/Residues: 1-107, 'S', 109-238 <PRA1>

A/Cross-references: GB:M62654; NID:G155662; PIDN:AAA27722.1; PID:G155663

A/Accession: JQ1514

A/Molecule type: mRNA

A/Residues: 1-99, 'P', 101-140, 'L', 142-218, 'V', 220-238 <PRA2>

A/Cross-references: GB:M62653; NID:G155660; PIDN:AAA27721.1; PID:G155661

A/Accession: P00335

A/Molecule type: Protein

A/Residues: 46-64/74-122/132-151/154-183/185-200 <PRA3>

R/Inouye, S.; Teufel, F.I.
FEBS Lett. 351, 211-214, 1994

A/Title: Evidence for redox forms of the Aequorea green fluorescent protein.

A/Reference number: S48693; MUID:94364470; PMID:8082767

A/Accession: S48693

A/Status: preliminary

A/Residues: 1-24, 'Q', 26-156, 'P', 158-171, 'K', 173-238 <INO>

A/Cross-references: GB:I29345; NID:9606383; PIDN:AAA58246.1; PID:9606384

submitted to the EMBL Data Library, January 1995

A/Reference number: S51330

A/Accession: S51330

A/Molecule type: mRNA

A/Residues: 1-13, 'V', 15-24, 'Q', 26-44, 'N', 46-153, 'G', 155-156, 'P', 158-171, 'K', 173-227, 'R', 229-238 <INO>

A/Cross-references: EMBL:X83959; NID:9634008; PIDN:CAA58789.1; PID:9634009

A/Experimental source: clone gfp2

A/Accession: S51331

A/Molecule type: mRNA

A/Residues: 1-24, 'Q', 26-29, 'R', 31-63, 'L', 85-153, 'G', 155-156, 'P', 158-171, 'K', 173-208, 'Q', 209-238 <INO>

A/Cross-references: EMBL:X83960; NID:9634010; PIDN:CAA58790.1; PID:9634011

A/Experimental source: clone gfp2

R/Yang, F.; Moss, L.G.; Phillips Jr., G.N.

submitted to the Brookhaven Protein Data Bank, August 1996

A/Reference number: A65692; PDB:1GFI

A/Contents: annotation; X-ray crystallography, 1.9 angstroms, residues 'A', 2-79, 'R', 81-95

A/Note: engineered sequence based on JQ1514, cloned and expressed in Escherichia coli

R/Yang, F.; Moss, L.G.; Phillips Jr., G.N.

Nat. Biotechnol. 14, 1246-1251, 1996

A/Title: The molecular structure of green fluorescent protein.

A/Reference number: A58953; MUID:96294543; PMID:9631087

A:Contents: annotation; X-ray crystallography, 1.9 angstroms
 C:Comment: This protein is excited by the photoprotein aequorin (see PIR:AQFNU) emitting
 C:Comment: The chromophore of this protein is formed by modification of Ser-dehydro-Tyr-
 C:Genetics:
 A:Gene: GFP
 A:Insertions: 69/3; 167/3
 C:Superfamily: green-fluorescent protein
 C:Keywords: chromoprotein; luminescence
 F/65-67/Cross-link: 5-imidazolone (Ser-Gly) #status experimental
 F/66/Modified site: dehydrotyrosine (Tyr) #status experimental

Query Match 83.0%; Score 1245; DB 1; Length 238;
 Best Local Similarity 97.1%; Pred. No. 1.5e-93;
 Matches 231; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 2 VSKSEELFTGVVPLIVELDGVNCHKFSVSGEGSDATYGLTLKFTCTTGKLPVPPPTL 61
 Db 1 MSKEELFTGVVPLIVELDGVNCHKFSVSGEGSDATYGLTLKFTCTTGKLPVPPPTL 60
 QY 62 VTTLTYGVQCFRYPDMKQHPFKSAMPRGYVQERTTFKDDGNYTRARVKEGDTLV 121
 Db 61 VTTFYSYGVQCFRYPDMKQHPFKSAMPRGYVQERTTFKDDGNYTRARVKEGDTLV 120
 QY 122 NRLEIKDIDKEDGNITLGHLELYNNHNYIMADKQNGIKYFKIRNIEDGSVOLAD 161
 Db 121 NRLEIKDIDKEDGNITLGHLELYNNHNYIMADKQNGIKYFKIRNIEDGSVOLAD 160
 QY 182 HYQONTPIGDGVPLLPNNHYLSTQSALSKDPEKEDMVLLEFTAGITLGMDELYK 239
 Db 181 HYQONTPIGDGVPLLPNNHYLSTQSALSKDPEKEDMVLLEFTAGITLGMDELYK 238

RESULT 2
 DCM50
 Ornithine decarboxylase (EC 4.1.1.17) [validated] - mouse
 C:Species: Mus musculus (house mouse)
 C>Date: 27-Nov-1985 #sequence revision 27-Nov-1985 #text change 18-Aug-2000
 C:Accession: A01077; A22295; S02124; A25427; B25427; I55283; S34065
 R:Kahana, C.; Nathans, D.
 Proc. Natl. Acad. Sci. U.S.A. 82, 1673-1677, 1985
 A:Title: Nucleotide sequence of murine ornithine decarboxylase mRNA.
 A:Reference number: A01077; MUID:8516193; PMID:3856848
 A:Accession: A01077

A:Molecule type: mRNA
 A:Residues: 1-461 <RAH>
 A:Cross-references: GB:M10624; NID:g200119; PIDN:AAA39845.1; PID:g200120
 R:Gupta, M.; Coffino, P.
 J. Biol. Chem. 260, 2941-2944, 1985
 A:Title: Mouse ornithine decarboxylase. Complete amino acid sequence deduced from cDNA.
 A:Reference number: A22295; MUID:85131068; PMID:2982844
 A:Accession: A22295
 A:Molecule type: mRNA
 A:Residues: 1-461 <GUP>
 A:Cross-references: GB:M19521; GB:M20617; GB:J03615; NID:g200127; PIDN:AAA51638.1; PID:g200124
 R:Coffino, P.; Chen, E. L.
 Nucleic Acids Res. 16, 2731-2732, 1988
 A:Title: Nucleotide sequence of the mouse ornithine decarboxylase gene.
 A:Reference number: S02124; MUID:88203223; PMID:3362685
 A:Accession: S02124
 A:Molecule type: DNA
 A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-461 <COF>
 A:Cross-references: EMBL:X07392; NID:g53517; PIDN:CAA30301.1; PID:g53518
 R:Richock, N.J.; Seppanen, P.T.; Kontula, K.K.; Janne, P.A.; Bardin, C.W.; Janne, O.A.
 Proc. Natl. Acad. Sci. U.S.A. 83, 594-598, 1986
 A:Title: Two ornithine decarboxylase mRNA species in mouse kidney arise from size hetero
 A:Reference number: A94117; MUID:86120987; PMID:3456155
 A:Accession: A25427
 A:Molecule type: mRNA
 A:Residues: 55-56,180-205, 'E', 207-461 <HIC1>
 A:Cross-references: GB:M12330; NID:g200121; PIDN:AAA9846.1; PID:g387499
 A:Note: this mRNA was designated as clone podc16
 A:Accession: B25427

A:Molecule type: mRNA
 A:Residues: 320-349, 'H', 351-461 <HIC2>
 A:Cross-references: GB:M12331; NID:g200130; PIDN:AAA9848.1; PID:g200131
 A:Note: this mRNA was designated as clone podc74
 A:Note: the authors translated the codon CAT for residue 330 as Tyr
 J.Poulin, R.; Lu, L.; Ackermann, B.; Bey, P.; Pegg, A.E.
 J. Biol. Chem. 267, 150-158, 1992
 A:Title: Mechanism of the irreversible inactivation of mouse ornithine decarboxylase by
 A:Reference number: A41744; MUID:92112641; PMID:1730582
 A:Contents: annotation; active site
 R:Katz, A.; Kahana, C.
 J. Biol. Chem. 263, 7604-7609, 1988
 A:Title: Isolation and characterization of the mouse ornithine decarboxylase gene.
 A:Reference number: I55283; MUID:88227956; PMID:3372502
 A:Accession: I55283
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-461 <RES>

A:Cross-references: GB:J03733; NID:g200132; PIDN:AAA9849.1; PID:g387501
 R:Teirka, S.R.; Turk, C.W.; Coffino, P.
 Biochem. J. 293, 289-295, 1993
 A:Title: Multiple active conformers of mouse ornithine decarboxylase.
 A:Reference number: S34065; MUID:93319524; PMID:8328969
 A:Accession: S34065
 A:Molecule type: protein
 A:Residues: 357-359, 'X', 361-367 <TS1>
 C:Genetics:
 A:Insertions: 34/3; 92/3; 150/2; 195/2; 222/3; 250/3; 305/1; 342/3; 414/2
 A:Complex: homodimer
 C:Function:

A:Description: catalyzes the decarboxylation of ornithine to putrescine
 A:Pathway: polyamine biosynthesis
 A:Note: the first and rate-limiting reaction in the pathway
 C:Superfamily: ornithine decarboxylase
 C:Keywords: carbon-carbon lyase; carboxy-lyase; homodimer; phosphoprotein; polyamine bio
 F/69/Binding site: pyridoxal phosphate (Lys) (covalent) #status experimental
 F/360/Active site: Cys (shared with dimeric partner) #status experimental

Query Match 14.8%; Score 221.5; DB 1; Length 461;
 Best Local Similarity 87.8%; Pred. No. 3.2e-10;
 Matches 43; Conservative 3; Mismatches 2; Indels 1; Gaps 1;

QY 234 MDELYKTL-SHGPFPEVEBODDGTLPMSCAQSGMDRHPAASARINV 281
 Db 413 MMQLMKQIQSHGFPEVEBODDGTLPMSCAQSGMDRHPAASARINV 461

RESULT 3
 156477
 Ornithine decarboxylase (EC 4.1.1.17) - mouse
 C:Species: Mus musculus domesticus (western European house mouse)
 C>Date: 19-Mar-1997 #sequence revision 19-Mar-1997 #text change 08-Sep-2000
 C:Accession: I56477
 R:Joannes, G.U.; Berger, F.G.
 J. Mol. Evol. 36, 555-567, 1993
 A:Title: Domains within the mammalian ornithine decarboxylase messenger RNA have evolved
 A:Reference number: I56477; MUID:93353527; PMID:8350350
 A:Accession: I56477
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-461 <JOH>
 A:Cross-references: GB:S64539; NID:g404561; PIDN:AB27809.1; PID:g404562
 C:Superfamily: ornithine decarboxylase
 C:Keywords: carbon-carbon lyase; carboxy-lyase; phosphoprotein; polyamine biosynthesis;
 F/69/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match 14.8%; Score 221.5; DB 2; Length 461;
 Best Local Similarity 87.8%; Pred. No. 3.2e-10;
 Matches 43; Conservative 3; Mismatches 2; Indels 1; Gaps 1;

QY 234 MDELYKTL-SHGPFPEVEBODDGTLPMSCAQSGMDRHPAASARINV 281
 Db 413 MMQLMKQIQSHGFPEVEBODDGTLPMSCAQSGMDRHPAASARINV 461

RESULT 4
I55356
O:Ornithine decarboxylase (EC 4.1.1.17) - shrew mouse
C:Species: Mus pahari
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 08-Sep-2000
A:Accession: I55356
R:Johannes, G.; Berger, F.G.
J. Biol. Chem. 267, 10108-10115, 1992
A:Title: Alterations in mRNA translation as a mechanism for the modification of enzyme
A:Reference number: I55356; MUID:92250571; PMID:1374399
A:Accession: I55356
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-461 <RES>
A:Cross-references: GB:M87223; NID:g200123; PIDN:AAA39847.1; PID:g200124
C:Genetics:
A:Gene: ODC
C:Superfamily: ornithine decarboxylase
C:Keywords: carbon-carbon lyase; carboxy-lyase; phosphoprotein; polyamine biosynthesis;
F:69/Binding site: pyridoxal phosphate (lys) (covalent) #status predicted
Query Match 14.0%; Score 210.5; DB 2; Length 461;
Best Local Similarity 29.1%; Pred. No. 2.5e-09;
Matches 78; Conservative 33; Mismatches 70; Indels 87; Gaps 12;
QY 90 PEGYVQ---ERTIFKDDGNKYTRAIVEEGDTLVNRIELKG-----IDFKEDGN 136
DB 205 PETFVQAVSDARCVF--DWG-----TEVGFS---MYLIDIGGPGSGEDTKIKFEETIS 253
QY 137 ILGHKLE-YNNVSNVYIMADKOKNGIKVNFKIRNI-----EDGSYQL 179
DB 254 VINPALDKYFPDSGVRILAEGRYVASFILAVNIIAKKTWKEQGSDEDESNQET 313
QY 180 ADHYQNTPTIGDGPVLLPDNHYLSTQSALSKDPNEK-----RDHNV----- 220
DB 314 EMYVNDGVYGSFNCILVDHAHVAKALLQKRPDPKESYSSISWGPTCDGLDRIVERCSLP 373
QY 221 -----LLE---FVTAAGITL-----MDLYKKL-SHGFPPEVEROD 253
DB 374 EMEVGDMMFLFENMGAYTVAASTFNGFQRPNTIYVMSRMWQLMKRIQSHGFPPEVEROD 433
QY 254 DGTLPMSCAQESGMDRHPAACASARINV 281
DB 434 DGTLPMSCAQESGMDHSHAAACASARINV 461
RESULT 5
DCRTO
O:Ornithine decarboxylase (EC 4.1.1.17) - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 18-Jun-1999
A:Accession: A27361; A33710; C33710; S01406; S14352
R:van Kranen, H.J.; van de Zande, L.; van Kreijl, C.F.; Bisschop, A.; Wieringa, B.
Gene 60, 145-155, 1987
A:Title: Cloning and nucleotide sequence of rat ornithine decarboxylase cDNA.
A:Reference number: A27361; MUID:88167817; PMID:3443298
A:Accession: A27361
A:Molecule type: mRNA
A:Residues: 1-461 <VAN>
A:Cross-references: GB:M16982; NID:g205803; PIDN:AAA41737.1; PID:g205804; GB:M19157
A:Experimental source: Clone pODC.F10
J. Mol. Biol. Chem. 264, 9016-9021, 1989
R:Wen, L.; Huang, J.K.; Blackshear, P.J.
A:Title: Rat ornithine decarboxylase gene. Nucleotide sequence, potential regulatory ele
A:Reference number: A33710; MUID:89255378; PMID:2722815
A:Accession: A33710
A:Molecule type: DNA
A:Residues: 1-461 <MEN1>
A:Cross-references: GB:J04792
A:Accession: C33710
A:Molecule type: mRNA

A:Residues: 1-461 <MEN2>
A:Cross-references: GB:J04791; NID:g205807; PIDN:AAA6164.1; PID:g205808
R:van Steeg, H.; van Oostrom, C.T.M.; van Kranen, H.J.; van Kreijl, C.F.
Nucleic Acids Res. 16, 8173-8174, 1988
A:Title: Nucleotide sequence of the rat ornithine decarboxylase gene.
A:Reference number: S01406; MUID:88335556; PMID:3419906
A:Accession: S01406
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-461 <VAN2>
A:Cross-references: EMBL:X07944; NID:g56785; PIDN:CAA30765.1; PID:g56787
R:van Steeg, H.; van Oostrom, C.T.M.; Hodemaekers, H.M.; Peters, L.; Thomas, A.A.M.
Biochem. J. 274, 521-526, 1991
A:Title: The translation in vitro of rat ornithine decarboxylase mRNA is blocked by its
A:Reference number: S14352; MUID:91174765; PMID:2006916
A:Accession: S14352
A:Molecule type: mRNA
A:Residues: 1-37 <STE>
A:Comment: This pyridoxal phosphate enzyme, which belongs to a multigene family, catalyz
in the pathway of polyamine biosynthesis.
C:Genetics:
A:Insertions: 34/3; 92/3; 150/2; 195/2; 222/3; 250/3; 305/1; 342/3; 414/2
C:Superfamily: Ornithine decarboxylase
C:Keywords: carbon-carbon lyase; carboxy-lyase; phosphoprotein; polyamine biosynthesis;
F:69/Binding site: pyridoxal phosphate (lys) (covalent) #status predicted
F:160/Active site: Cys (shared with dimeric partner) #status predicted
Query Match 13.9%; Score 208.5; DB 1; Length 461;
Best Local Similarity 29.1%; Pred. No. 3.6e-09;
Matches 78; Conservative 33; Mismatches 70; Indels 87; Gaps 12;
QY 90 PEGYVQ---ERTIFKDDGNKYTRAIVEEGDTLVNRIELKG-----IDFKEDGN 136
DB 205 PETFVQAVSDARCVF--DWG-----TEVGFS---MYLIDIGGPGSGEDTKIKFEETIS 253
QY 137 ILGHKLE-YNNVSNVYIMADKOKNGIKVNFKIRNI-----EDGSYQL 179
DB 254 VINPALDKYFPDSGVRILAEGRYVASFILAVNIIAKKTWKEQGSDEDESNQET 313
QY 180 ADHYQNTPTIGDGPVLLPDNHYLSTQSALSKDPNEK-----RDHNV----- 220
DB 314 EMYVNDGVYGSFNCILVDHAHVAKALLQKRPDPKESYSSISWGPTCDGLDRIVERCSLP 373
QY 221 -----LLE---FVTAAGITL-----GMDELYKKL-SHGFPPEVEROD 253
DB 374 EMEVGDMMFLFENMGAYTVAASTFNGFQRPNTIYVMSRMWQLMKRIQSHGFPPEVEROD 433
QY 254 DGTLPMSCAQESGMDRHPAACASARINV 281
DB 434 VGTLPMSCAQESGMDRHPAACASARINV 461
RESULT 6
DCHYC
O:Ornithine decarboxylase (EC 4.1.1.17) - Chinese hamster
C:Species: Cricetus griseus (Chinese hamster)
C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 18-Jun-1999
A:Accession: S09574; A27379
R:Grens, A.; Steglich, C.; Pilz, R.; Scheffler, I.E.
Nucleic Acids Res. 17, 10497, 1989
A:Title: Nucleotide sequence of the Chinese hamster ornithine decarboxylase gene.
A:Reference number: S09574; MUID:90098890; PMID:2602162
A:Accession: S09574
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-455 <GEB>
A:Cross-references: EMBL:X16910; NID:g49439; PIDN:CAA4784.1; PID:g49440
R:Srinivasan, P.R.; Tonin, P.N.; Wensing, E.J.; Lewis, W.H.
J. Biol. Chem. 262, 12871-12878, 1987
A:Title: The gene for ornithine decarboxylase is co-amplified in hydroxyurea-resistant he
A:Reference number: A27379; MUID:87308323; PMID:2887574
A:Accession: A27379
A:Molecule type: mRNA

A:Residues: 163-295, 'EOP', 296-451, 'R', 453-455 <SRI>
 A:Cross-references: GB:U02813
 C:Superfamily: ornithine decarboxylase
 C:Keywords: carbon-carbon lyase; carboxy-lyase; phosphoprotein; polyamine biosynthesis;
 F:67/Binding site: pyridoxal phosphate (lys) (covalent) #status predicted
 F:354/Active site: Cys (shared with dimeric partner) #status predicted

Query Match 13.3%; Score 200; DB 1; Length 455;
 Best Local Similarity 28.7%; Pred. No. 1,7e-08;
 Matches 76; Conservative 34; Mismatches 71; Indels 84; Gaps 12;

QY 90 PEGYVQ-----ERTIFPKDQNGVKTAEVKEGTLVNRILKQ-----IDFKEDGN 136
 Db 202 PETFVQALSDARCVF---DMG-----TEVGS-----WTLDIGGFGPGESEDTKLKFEITIS 250
 QY 137 ILGHKLE-YNYNHNVYIMADKQNGIKVNFKIRHN-----EDGSYQLADH 182
 Db 251 VINPALDKYPPPSGKRVIAEPGRYYASAFTLAVNIIAKKIVSKSGSDDESSSQTFMY 310
 QY 183 YQONTPIGDGVPVLLPDNHYLSTQSLSKDPNEK-----RDHNV----- 220
 Db 311 YVNDGVYGSFNCILYDHAHVKPLLPKRPKPEKYSISIWGPTCDGLDRIVERCNLPENH 370
 QY 221 -----LLE-----FVTAAGITLG-----MDELYKTL-SHGPPPEVEEDDGT 256
 Db 371 VGDWMLFENMGAYTVAAASTFNCFORPSIYYVMSRPMWQIMKQIQNHGPFPEVEEDVGT 430
 QY 257 LPMSCAQSQSGMDRHPAACAASRINV 281
 Db 431 LPTSCAQSGQMDRHPAACAASINIV 455

RESULT 7
 DCHDO
 ornithine decarboxylase (EC 4.1.1.17) - human
 N:Alternate names: ODC
 C:Species: Homo sapiens (man)
 C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 18-Jun-1999
 C:Accession: S06900; A33388; J00792; A26957; S14539; A46555; I38077
 R:Van Steeg, H.; van Oostrom, C.T.M.; Martens, J.W.M.; van Kreyl, C.F.; Schepens, J.; W
 Nucleic Acids Res. 17, 8855-8856, 1989
 A:Title: Nucleotide sequence of the human ornithine decarboxylase gene.
 A:Reference number: S06900; MUID:90067851; PMID:2587220
 A:Accession: S06900
 A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-461 <VAN>
 A:Cross-references: EMBL:X16277; NID:g35137; PIDN:CAA34353.1; PID:g236667
 R:Flitzgerald, M.C.; Flanagan, M.A.
 DNA 8, 623-634, 1989
 A:Title: Characterization and sequence analysis of the human ornithine decarboxylase gen
 A:Reference number: A33388; MUID:90126232; PMID:2693021
 A:Accession: A33388
 A:Molecule type: DNA
 A:Residues: 1-461 <FIT>
 A:Cross-references: GB:M01740; NID:g189370; PIDN:AAAS967.1; PID:g189371; GB:M01061; NID
 R:Hiickok, N.J.; Seppanen, P.J.; Gunsalus, G.L.; Jaenne, O.A.
 DNA 6, 179-187, 1987
 A:Title: Complete amino acid sequence of human ornithine decarboxylase deduced from comp
 A:Reference number: A26957; MUID:87246067; PMID:3595418
 A:Accession: A26957
 A:Molecule type: mRNA
 A:Residues: 1-461 <HIC2>
 A:Cross-references: GB:M16650
 R:Flanagan, M.A.; Streng, K.A.; Wagner III, R.L.

submitted to the EMBL Data Library, November 1990
 A:Description: Nucleotide sequence of a human ornithine decarboxylase cDNA.
 A:Reference number: S14539
 A:Accession: S14539
 A:Molecule type: mRNA
 A:Residues: 1-461 <FLA>
 A:Cross-references: EMBL:X55362; NID:g35135; PIDN:CAA39047.1; PID:g35136
 R:Kaczmarek, L.; Calabretta, B.; Ferrati, S.; de Riel, J.K.
 J. Cell. Physiol. 132, 545-551, 1987
 A:Title: Cell-cycle-dependent expression of human ornithine decarboxylase.
 A:Reference number: A46555; MUID:88007889; PMID:3308908
 A:Accession: A46555
 A:Molecule type: mRNA
 A:Residues: 112-461 <KAC>
 A:Cross-references: GB:M20372; NID:g189372; PIDN:AAAS968.1; PID:g189373
 R:Heien, U.T.; Deming, M.F.; Heide, S.M.; Verma, A.K.
 Cancer Res. 50, 2239-2244, 1990
 A:Title: Expression of human chromosome 2 ornithine decarboxylase gene in ornithine decar
 A:Reference number: I38077; MUID:90199754; PMID:2317811
 A:Accession: I38077
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-22 <HS1>
 A:Cross-references: EMBL:X53271; NID:g288102; PIDN:CAA37369.1; PID:g288103
 C:Genetics:
 A:Gene: GDB:ODC1
 A:Cross-references: GDB:119462; OMIM:165640
 A:Map position: 2p25-2p25
 A:Intons: 34/3; 92/3; 150/2; 195/2; 222/3; 250/3; 305/1; 342/3; 414/2
 C:Function:
 A:Description: catalyzes the decarboxylation of ornithine to putrescine
 A:Pathway: polyamine biosynthesis
 A>Note: the first and rate-limiting reaction in the pathway
 C:Superfamily: ornithine decarboxylase
 C:Keywords: carbon-carbon lyase; carboxy-lyase; phosphoprotein; polyamine biosynthesis;
 F:69/Binding site: pyridoxal phosphate (lys) (covalent) #status predicted
 F:360/Active site: Cys (shared with dimeric partner) #status predicted

Query Match 10.3%; Score 154.5; DB 1; Length 461;
 Best Local Similarity 27.2%; Pred. No. 8.7e-05;
 Matches 72; Conservative 29; Mismatches 83; Indels 81; Gaps 12;

QY 90 PEGYVQ-----ERTIFPKDQNGVKTAEVKE-----GDTLVNRILEKGIDPKEDGNILG 139
 Db 205 PETFVQALSDARCVF---DMG-----AYGFSMYLIDIGGCGPPGSDVA-LKFEETGVIN 256
 QY 140 HKLE-YNYNHNVYIMADKQNGIKVNFKIRHN-----EDGSYQLADH 182
 Db 257 PALDKYPPPSDGVRIIAEPGRYYASAFTLAVNIIAKKIVLKEQTSDDDESSSQTFMY 316
 QY 183 YQONTPIGDGVPVLLPDNHYLSTQSLSKDPNEK-----RDHNV----- 220
 Db 317 YVNDGVYGSFNCILYDHAHVKPLLPKRPKPEKYSISIWGPTCDGLDRIVERCDLPENH 376
 QY 221 -----LLE-----FVTAAGITLG-----MDELYKTL-SHGPPPEVEEDDGT 256
 Db 377 VGDWMLFENMGAYTVAAASTFNCFORPSIYYVMSRPMWQIMKQIQNHGPFPEVEEDVGT 436
 QY 257 LPMSCAQSQSGMDRHPAACAASRINV 281
 Db 437 LPVSCAQSQSGMDRHPAACAASINIV 461

RESULT 8
 DCHDO
 ornithine decarboxylase (EC 4.1.1.17) - chicken (fragment)
 C:Species: Gallus gallus (chicken)
 C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 18-Jun-1999
 C:Accession: A48386; S19892
 R:Johnson, R.; Bulfield, G.
 Anim. Genet. 23, 403-409, 1992
 A:Title: Molecular cloning and sequence analysis of a chicken ornithine decarboxylase cDN
 A:Reference number: A48386; MUID:93036582; PMID:1416246

A:Accession: A48386
 A:Molecule type: mRNA
 A:Residues: 1-450 <002>
 A:Cross-references: EMBL:X64710; NID:963712; PIDN:CAA5965.1; PID:963713
 A:Experimental source: embryos
 A:Note: sequence extracted from NCBI backbone (NCBI:P115105)
 C:Function:
 A:Description: This enzyme catalyzes the decarboxylation of ornithine to putrescine, the C:Superfamily: ornithine decarboxylase
 C:Keywords: carbon-carbon lyase; carboxy-lyase; phosphoprotein; polyamine biosynthesis;
 F:5/Binding site: pyridoxal phosphate (lys) (covalent) #status predicted
 F:350/Active site: Cys (shared with dimeric partner) #status predicted

Query Match 7.4%; Score 111.5; DB 1; Length 450;
 Best Local Similarity 64.9%; Pred. No. 0.26;
 Matches 24; Conservative 5; Mismatches 7; Indels 1; Gaps 1;

Qy 245 PPEVEEDDGLTLPMSQSGMDRHPACASARINV 281
 Db 415 FLAVEEDQVASLPLSCAESGIE-YPAFCASASINV 450

RESULT 9

A:Accession: A43563
 A:Molecule type: mRNA
 A:Residues: 1-460 <BAS>
 A:Cross-references: GB:X56316; NID:964953; PIDN:CAA39760.1; PID:964954
 C:Superfamily: ornithine decarboxylase
 C:Keywords: carbon-carbon lyase; carboxy-lyase; phosphoprotein; polyamine biosynthesis;
 F:69/Binding site: pyridoxal phosphate (lys) (covalent) #status predicted

Query Match 6.9%; Score 103; DB 2; Length 460;
 Best Local Similarity 23.8%; Pred. No. 1.3;
 Matches 64; Conservative 34; Mismatches 81; Indels 90; Gaps 14;

Qy 90 PEGVVO---BRTFFKDDGNYKTRAEVFEEDTLVNRLEKG-----IDPKEDGN 136
 Db 205 PQTVOAVSDARCVF---DMG-----AELGFN---MHLLDGGGPPGSEDEVKAKREELTS 253
 Qy 137 ILGHKLE-YNNNSHNYITMAKOKNGIKVNF-----KIRHNIEDGSVOLAD----- 181
 Db 254 VINPALDYPADSGVKLIAEPGRYYVASPTLAVNITAKIVWVEQSGSDDEEDANDK 313
 Qy 182 ---HYQONTPTGIDGVLLEPDNHYLSTGSLSKDPNEK-----RDHNV----- 220
 Db 314 TLMYYVDGVGSGNCLIFDAHAKPVITKPKPEKPEKSSSWGPGTCDGIDRIVEREL 373
 Qy 221 -----ILE-----FVTAAGITLG---MDLYKKLS-----HGFPPEVEEQ 252
 Db 374 PELQGVDMWLFENMGAYVAALSTFNGFRPLTYVSRPMQIMHDIKKGILPEV--P 431
 Qy 253 DDGTLPMSCQSGMDRHPACASARINV 281
 Db 432 DLSALHVSQAGSGMELAPAVCTASINV 460

RESULT 10

A:Accession: H72228
 A:Molecule type: protein
 A:Residues: 1-253
 A:Cross-references: EMBL:X64710; NID:963712; PIDN:CAA5965.1; PID:963713
 A:Experimental source: embryos
 A:Note: sequence extracted from NCBI backbone (NCBI:P115105)
 C:Function:
 A:Description: This enzyme catalyzes the decarboxylation of ornithine to putrescine, the C:Superfamily: ornithine decarboxylase
 C:Keywords: carbon-carbon lyase; carboxy-lyase; phosphoprotein; polyamine biosynthesis;
 F:5/Binding site: pyridoxal phosphate (lys) (covalent) #status predicted
 F:350/Active site: Cys (shared with dimeric partner) #status predicted

Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.; C.M.
 Nature 399, 323-329, 1999
 A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
 A:Reference number: A12200; MUID:99287316; PMID:10360571
 A:Accession: H72228
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-785 <ARN>
 A:Cross-references: GB:AE001806; GB:AE000512; NID:94982196; PIDN:AA036691.1; PID:9498219
 A:Experimental source: strain MSB8
 C:Genetics:
 A:Gene: TM1624

Query Match 6.7%; Score 101; DB 2; Length 785;
 Best Local Similarity 20.9%; Pred. No. 3.8;
 Matches 40; Conservative 27; Mismatches 64; Indels 60; Gaps 5;

Qy 3 SKGERLFTGVVPIVELDGVNGHKFSVSGEGDATYGLTIKFTCTGKLPVMPITLV 62
 Db 15 NEGRSFEGTVPGVQAD-----LVKGLLPHRYGSM- 46
 Qy 63 TTLTYGVOCFSRRYPDHMKQHDFFKSAMEGVYQERTTFKDDGNYKTRAEVFEEDTLV 122
 Db 47 -----NEDLFKEIEDREWITREREFKEDVKEGERVDLVFEGVDTL 88
 Qy 123 RIELKIDPKEDGNILGHKLEYYNNNSHNYITMAKOKNGIKVNFKIRHNIEDGSVOLADH 182
 Db 89 DVLANGVTL---GSTEDMFIRPDVINVL---KEKNHLKVIK-----SPIRVEKT 134
 Qy 183 YQONTPTGIDGP 193
 Db 135 LEQVYVGLGP 145

RESULT 11

ERH04H
 coatomer complex alpha chain homolog - human
 A:Alternate names: HEP-COP; xenopsin homolog; xenopsin-related peptide precursor
 N:Contents: xenin 25; xenopsin-related peptide

C:Species: Homo sapiens (man)
 C:Date: 10-May-1996 #sequence
 C:Accession: J04668; A44317
 R:Chow, V.T.K.; Quek, H.H.
 Gene 169, 223-227, 1996

A:Title: HEP-COP, a novel human gene whose product is highly homologous to the alpha-sub
 A:Reference number: J04668; MUID:96194806; PMID:8647451
 A:Accession: J04668

A:Molecule type: mRNA
 A:Residues: 1-1224 <CHO>
 A:Cross-references: GB:U24105; NID:91638873; PIDN:AA070879.1; PID:91002369
 A:Experimental source: Hep3b hepatocellular carcinoma cell
 J:Feurle, G.E.; Hamscher, G.; Kusiek, R.; Meyer, H.E.; Metzger, J.W.
 J. Biol. Chem. 267, 22305-22309, 1992
 A:Title: Identification of xenin, a xenopsin-related peptide, in the human gastric mucose
 A:Reference number: A44317; MUID:93054515; PMID:1429581
 A:Accession: A44317
 A:Molecule type: protein
 A:Residues: 1-25 <FEU>
 A:Experimental source: gastric mucosa
 A:Note: sequence extracted from NCBI backbone (NCBI:P117018)
 A:Note: plasma levels of xenin 25, as determined by immunoassay, rise after meals and it
 C:Genetics:
 A:Gene: GDB:COA: HEP-COP

A:Cross-references: GDB:4642787; OMIM:601924
 A:Map position: 1q23-1q25
 C:Superfamily: yeast coatomer complex alpha chain; WD repeat homology
 C:Keywords: duplication; hormone; plasma; stomach
 F:1-25/Product: xenin 25 #status experimental <XNP>
 F:5-38/Domain: WD repeat homology <WD1>
 F:17-25/Product: xenopsin-related peptide #status predicted <XRP>
 F:47-80/Domain: WD repeat homology <WD2>
 F:89-122/Domain: WD repeat homology <WD3>

F:131-164/Domain: WD repeat homology <WD4>
F:201-234/Domain: WD repeat homology <WD5>
F:245-278/Domain: WD repeat homology <WD6>

Query Match 6.6%; Score 99.5; DB 1; Length 1224;
Best Local Similarity 26.1%; Pred. No. 9.2;

Matches 47; Conservative 28; Mismatches 68; Indels 37; Gaps 8;

QY 91 EGVYQERTFFPKDGNKYTRAVEREGDTLVNRLE-LKGDIFKEDGILGKLEYNVNSH 149

Db 632 KGYEVALHFKDE---KTRFSLALECGNIEIALMAKALDDKNCWEKLEVALLOQNHQ 688

QY 150 NYVIMADKQKNGIKVNV-----XIRNIEDGSVQ--LADHYQNTPIGDGPVLLPD 198

Db 689 IYKCYQRTNKPDPKSPFLYLITGNLEKTRKMKIAIRKMSGHYQVALYLD----- 741

QY 199 NHYLSQSLSKDPNEKRDHNVLEFVTAAGITLGMDELKXLSHGPPEVVEKDDGTLF 258

Db 742 ---VSEKRVRLIKKCGCKS-----LAYITLA--THGLDEAESLKETFPDPKE-----TTP 786

RESULT 12

T20721
hypothetical protein F25F2.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 19-May-2000

C:Accession: T20721; T21343; T23842

R:Ainscough, R.

submitted to the EMBL Data Library, August 1994

A:Reference number: Z19314

A:Accession: T20721

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-4307 <MT2>

A:Cross-references: EMBL:Z35599; PIDN:CAA84721.1; GSPDB:GN00021; CESP:F25F2.2

A:Experimental source: clone F10G11

R:Ainscough, R.

submitted to the EMBL Data Library, August 1994

A:Reference number: Z19410

A:Accession: T21343

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-4307 <MT2>

A:Cross-references: EMBL:Z35599; PIDN:CAA84661.1; GSPDB:GN00021; CESP:F25F2.2

A:Experimental source: clone F25F2

R:Ainscough, R.

submitted to the EMBL Data Library, June 1994

A:Reference number: Z19806

A:Accession: T23842

A:Status: preliminary; translated from GB/EMBL/DBJ

QY 101 FKD-----DGNVKT-RAVKEFGDTLVNRLELKGDIFKEDGILGKLEYNV 146

Db 3433 DEDEHQNGILRFGVLEGSIOSEPRAOVQSRSTHLPFRVDPNTGDIWSD-----HSITQGL 3487

QY 147 NSHVYIMADKQKNGIK-----VNFKIR----- 169

Db 3488 HFEFNVTV-IDSKEFNVSVYEVHVTSIDNDVIDHAVISIRIRMSVDPEFMKHYKEPRRIS 3546

QY 170 HNI---EDGSVOLADHYQNTPIGDGPVLLPDNHYLSQSLSKDPNEKRDHNVLEFVT 226

Db 3547 HHNLNDDSSIQLI--VQAVP-----STES-----ERRSRNSEDVEILM 3586

QY 227 AAGITLGM-----DELKXLSHGF 245

Db 3587 TNAQGLGRTGLKPDHITSLKNDPF 3610

RESULT 13

KXRTCI

proprotein convertase 1 (EC 3.4.21.93) precursor - rat

N:Alternate names: furin homolog PC1; kexin homolog PC1; prohormone cleavage enzyme; prot

C:Species: Rattus norvegicus (Norway rat)

C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 19-May-2000

C:Accession: A41556; S27361; S36358

R:Bloomquist, B.T.; Hipper, B.A.; Maine, R.E.

Mol. Endocrinol. 5, 2014-2024, 1991

A:Reference number: A41556; MUID:92168040; PMID:1791845

A:Accession: A41556

A:Molecule type: mRNA

A:Residues: 1-752 <BLD>

A:Cross-references: GB:M76705; MID:G203508; PIDN:AAA40945.1; PID:G203509

R:Hakes, D.J.; Birch, N.P.; Mezey, A.; Dixon, J.E.

Endocrinology 129, 3053-3063, 1991

A:Title: Isolation of two complementary deoxyribonucleic acid clones from a rat insulinoma

to endocrine and neuroendocrine tissues in rats.

A:Reference number: S27361; MUID:92063860; PMID:1954888

A:Accession: S27361

A:Molecule type: mRNA

A:Residues: 1-513, 'A', 515-752 <HAK>

A:Cross-references: EMBL:M83745

R:Hakes, D.J.; Birch, N.P.; Mezey, A.; Dixon, J.E.

submitted to the EMBL Data Library, February 1992

A:Reference number: S36358

A:Accession: S36358

A:Molecule type: mRNA

A:Residues: 1-366, 'T', 367-513, 'A', 515-752 <HAN>

A:Cross-references: EMBL:M83745; MID:G205062; PIDN:AAA41476.1; PID:G205063

C:Comment: This protein is a member of a family of subtilisin-like proteinases responsibl

C:Comment: This protein lacks a classical hydrophobic transmembrane segment but may assoc

C:Keywords: glycoprotein; hydrolase; serine proteinase

F.1.24/Domain: signal sequence #status predicted <SIG>

F.25-110/Domain: propeptide #status predicted <PRO>

F.111-752/Product: prohormone-processing proteinase PC1 #status predicted <MAT>

F.158-396/Domain: subtilisin homology <SBT>

F.167-208,382/Active site: Asp, His, Ser #status predicted

F.173,401,645/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 6.2%; Score 92.5; DB 1; Length 752;

Best Local Similarity 24.4%; Pred. No. 18;

Matches 60; Conservative 29; Mismatches 80; Indels 77; Gaps 13;

QY 18 ELDDVNGHK-----FSVSGEGEDATGKTLKRICITGKLP-----VWPVTLVITLYTG 68

Db 540 ERLDSPNGFKWMDMSVHTVGENFV--GVTWLTWDSGRQNGRIVNKLII-----HG 593

QY 69 VQCFSRYPDNHKNQDFFKSNMPEGYVQERTIFFKDDGNKYTRAVEREGDTLVNRILEKG 128

Db 594 T---SSQPEHMKQ-----PRVYT-----SYMTVQNDRGRGVEKVVAVVEK 631

QY 129 IDRFEDGNILGKLEYNVNSHVYIMADKQKNGIKVNFKIRNIEDGSVOLADHYQNTPI 188

Db 632 TONSINGNLVPE--NSSSSSVEDRRDECVQCAPKAMIR-----LLQSAFSEKNTP 680
 QY 189 IGDGPVLLPDNHYLSTQSAISKDPNEKRDHMLLEFVTAAGITLGMDELTKLSHGFPPE 248
 Db 681 -----SKQS--SKTSPAKLS-----VPYEGLYEALFKLNK-----PSQ 711
 QY 249 VERQDD 254
 Db 712 LEDSED 717

RESULT 14

C64468
 hypotetrical protein M01348 - Methanococcus jannaschii
 CjDate: 13-Sep-1996 #sequence_revision 13-Sep-1996 #ext_change 21-Jul-2000
 CjAccession: C64468
 R.Bull, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,
 ; Reich, C.I.; Overbeek, R.; Kitzness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;
 rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
 Science 273, 1058-1073, 1996
 A:Authors: Kaine, B.P.; Borodovsky, M.; Kleen, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C
 A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii
 A:Reference number: A64300; M01D:96337993; PMID:8688087
 A:Accession: C64468
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-336 <BU>
 A:Cross-references: GB:U67574; GB:L77117; NID:g1591978; PIDN:AB99360.1; PID:g1591990; T
 C:Genetics:
 A:Map position: FOR1295121-1296131
 A:Start codon: TTG

Query Match 6.1%; Score 92; DB 2; Length 336;
 Best Local Similarity 24.9%; Pred. No. 6.6;
 Matches 66; Conservative 37; Mismatches 80; Indels 82; Gaps 17;
 QY 20 DGDVNGHFSVSGE---GEGDAYGKL--TKFTCTTGKLPVPW-----PTL 61
 Db 91 DGDY---YNSGELSTASIFAKIGKIDITGNFNSGGE---WYNDVIKDANSEDL 142
 QY 62 VTTLTYGV-----QCFERYPDHMQ-----HDFKSAPEGYQERTIFPKDDG 105
 Db 143 KSVLTFRDSYERKEILNRKP-HLKKLFEDNIYFNNSDFDFPM-----MEFIGAG 192
 QY 106 NYKTRAEVKEF-----GDTLVNRIELKSIDFKEDGNIL--GHKLEYNVNSHNYI 153
 Db 193 NMRKFLVEYEFKFKIKSCQISNEIYNEI-IRPD-KMSDLALAHLLKENYEKCLYV 250
 QY 154 MADK-----QKNGIVNFKIRHNIDGSVQLADHYQONTPIGDGPVLLP-DNH 201
 Db 251 MLFKEYFDLDEFNEIKKKIMLYDIAYNLKNGVKKKEWMLRNLKIKRPLPTY 310
 QY 202 LSTQSALSKD-PNEKRDHMLLEFY 225
 Db 311 ---KDAHNDLNLILDYVYVKEFI 332

RESULT 15

G81355
 tRNA (uracil-5-)-methyltransferase (EC 2.1.1.35) Cj0831c [imported] - Campylobacter jejuni
 CjSpecies: Campylobacter jejuni
 CjDate: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002
 CjAccession: G81355
 R.Parkhill, J.; Wren, B.W.; Mungall, K.; Kestley, J.M.; Churcher, C.; Basham, D.; Chilling,
 C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVlier, A.; Whitehead, S.; Barrell
 Nature 403, 665-668, 2000
 A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyf
 A:Reference number: A81250; M01D:20150912; PMID:10688204
 A:Accession: G81355
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-357 <PAR>

A:Cross-references: GB:AL139076; GB:AL111168; NID:96968128; PIDN:CA873096.1; PID:9696827
 A:Experimental source: serotype O2, strain NCTC 11168
 C:Genetics:
 A:Gene: trmA, Cj0831c
 C:Keywords: methyltransferase; S-adenosylmethionine

Query Match 6.1%; Score 91; DB 2; Length 357;
 Best Local Similarity 19.5%; Pred. No. 8.6;
 Matches 53; Conservative 43; Mismatches 84; Indels 92; Gaps 11;
 QY 80 KOHDFKSAPEGYQERTIFPKDGNKYTRAVERF--EGDTLV----- 121
 Db 14 EKHSPFKYKPEFFYKDPKLFASKDKKHRTAELSFYHNDTLFYAMDPSKKKXITLEY 73
 QY 122 -----NRIELF--GIDFKEDGNILGHKLEYNV----- 147
 Db 74 LDFADEKICAMPFRLLEVIARODNKKLEKLFGEVFFYTKQELSTLLYHKNLIEDIKSNLEN 133
 QY 148 -SHNYIMADQKNGIKYNEKIRH-----NIEDGSVQLADHYQONTPIGDGPVLLPDNHY 201
 Db 134 LSNLHNLILAKSKGKKLIFTEMLRQTLNTQDRKI---FYEFN-----NDCF 178
 QY 202 LSTQSALSKDPNEK-----RDHMLLEFVTAAG-ITLGMDELTKLSHGFPPEV 249
 Db 179 IQPNTAL---NEKMIWVCSILNTQKMDLLELYCGYGNFTLALAPFFKI---LATBI 231
 QY 250 EEODDGTLPMSCAQSGKMDRPPACAGARINV 281
 Db 232 SKSNINPALKNCELNNTTNIHFARLSSELSL 263

Search completed: May 14, 2004, 08:27:17
 Job time : 24 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 14, 2004, 08:23:58 ; Search time 17 Seconds
(without alignments)
860.689 Million cell updates/sec

Title: US-09-931-232-1

Perfect score: 1500

Sequence: 1 MYSKGEELFGVPIVLELD.....AQESGMDRHPAACAARINV 281

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1251	83.4	238	1	GFP_ABOVI
2	221.5	14.6	461	1	P00860 mus musculu
3	210.5	14.0	461	1	P27113 mus pahari
4	208.5	13.9	461	1	P09057 rattus norv
5	200	13.3	455	1	P14019 cricetus
6	154.5	10.3	461	1	DCOR_HUMAN
7	142.5	9.5	461	1	DCOR_HUMAN
8	111.5	7.4	450	1	DCOR_CHICK
9	103	6.9	460	1	DCOR_XENLA
10	99.5	6.6	1224	1	COPA_HUMAN
11	96.5	6.4	1224	1	COPA_BOVIN
12	92.5	6.2	752	1	NECI_RAT
13	92	6.1	336	1	YD48_METUA
14	91	6.1	357	1	TRMA_CAMLE
15	91	6.1	533	1	CP51_CANGA
16	90.5	6.0	861	1	SYL_HABIN
17	90.5	6.0	1193	1	RPOB_BACSU
18	89	5.9	439	1	SY62_DISOM
19	88.5	5.9	879	1	SYL_XYLP
20	88.5	5.9	866	1	ITHI_MESAU
21	88.5	5.9	6632	1	UN89_CAREL
22	87.5	5.8	658	1	ADAS_HUMAN
23	87	5.8	689	1	AC2L_HUMAN
24	86.5	5.8	461	1	PSBC_CYPAN
25	86.5	5.8	504	1	YC03_XLEPN
26	86.5	5.8	658	1	ADAS_CAYPO
27	86.5	5.8	860	1	SYL_SALT
28	86.5	5.8	860	1	SYL_SALT
29	86.5	5.8	860	1	SYL_SALT
30	86.5	5.8	1603	1	VIT4_CAREL
31	86.5	5.8	2222	1	DPOE_YEAST
32	85.5	5.7	788	1	DPOL_HPRHE
33	85.5	5.7	879	1	SYL_XYLET

34	85.5	5.7	1164	1	BAG_STRAG
35	85	5.7	466	1	Y065_MYCGE
36	85	5.7	3744	1	TRAI_YEAST
37	84.5	5.6	484	1	RGS9_BOVIN
38	84.5	5.6	613	1	PEPF_MYCPU
39	84.5	5.6	795	1	D152_HAERIN
40	84.5	5.6	797	1	D151_HAERIN
41	83.5	5.6	793	1	D153_HAERIN
42	83.5	5.6	859	1	SYL_SHEON
43	83	5.5	366	1	SET7_HUMAN
44	83	5.5	1214	1	BRF1_HUMAN
45	83	5.5	1585	1	PKX3_DICDI

ALIGNMENTS

RESULT 1	ID	GFP_ABOVI	STANDARD	PRT	238 AA.
AC	P42212	0171104	Q27903		
DT	01-NOV-1995	(Rel. 32, Created)			
DT	01-NOV-1995	(Rel. 32, Last sequence update)			
DT	10-OCT-2003	(Rel. 42, Last annotation update)			
DE	Green fluorescent protein.				
GN	GFP.				
OS	Aequorea victoria (Jellyfish).				
OC	Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Leptomedusae;				
OC	Aequoreidae; Aequorea.				
OX	NCBI_TaxID=6100;				
RN	[1]				
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.				
RX	MEDLINE=92175527; PubMed=1347277;				
RA	Prasher D.C., Eckhardt V.K., Ward W.W., Prendergast F.G.,				
RA	Cormier M.J.;				
RT	"Primary structure of the Aequorea victoria green-fluorescent				
RT	protein.";				
RL	Gene 111:229-233 (1992).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=94185810; PubMed=8137953;				
RA	Inouye S., Tsuji F.I.;				
RT	"Aequorea green fluorescent protein. Expression of the gene and				
RT	fluorescence characteristics of the recombinant protein.";				
RL	FEBS Lett. 341:277-280 (1994).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=9729832; PubMed=9154981;				
RA	Rouwendal G.J.A., Mendes O., Wolbert B.J.H., de Boer A.D.;				
RT	"Enhanced expression in tobacco of the gene encoding green fluorescent				
RT	protein by modification of its codon usage.";				
RL	Plant Mol. Biol. 33:989-999 (1997).				
RN	[4]				
RP	CHROMOPHORE.				
RX	MEDLINE=93192221; PubMed=8448132;				
RA	Cody C.W., Prasher D.C., Westler W.M., Prendergast F.G., Ward W.W.;				
RT	"Chemical structure of the hexapeptide chromophore of the Aequorea				
RT	green-fluorescent protein.";				
RL	Biochemistry 32:1212-1218 (1993).				
RN	[5]				
RP	X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).				
RX	MEDLINE=96355665; PubMed=8703075;				
RA	Omme M., Gubitt A.B., Kallio K., Gross L.A., Tsien R.Y.,				
RT	Remington S.J.;				
RT	"Crystal structure of the Aequorea victoria green fluorescent				
RT	protein.";				
RL	Science 273:1392-1395 (1996).				
RN	[6]				
RP	X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).				
RX	MEDLINE=98294543; PubMed=9631087;				
RA	Yang F., Moss L.G., Phillips G.N. Jr.;				
RT	"The molecular structure of green fluorescent protein.";				
RL	Nat. Biotechnol. 14:1246-1251 (1996).				

[7]
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF MUTANT WITH YELLOW EMISSION.
 RX MEDLINE=98455509; PubMed=9782051;
 RA Wachter R.M., Elsliiger M.A., Kallio K., Hanson G.T., Remington S.J.;
 RT "Structural basis of spectral shifts in the yellow-emission variants
 RT of green fluorescent protein.";
 RL Structure 6:1267-1277(1998).
 RN [8]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
 RX MEDLINE=99338303; PubMed=10220315;
 RA Elsliiger M.A., Wachter R.M., Hanson G.T., Kallio K., Remington S.J.;
 RT "Structural and spectral response of green fluorescent protein
 RT variants to changes in pH.";
 RL Biochemistry 38:5296-5301(1999).
 CC -1- FUNCTION: Energy-transfer acceptor. Its role is to transduce the
 CC blue chemiluminescence of the protein aequorin into green
 CC fluorescent light by energy transfer. Fluoresces in vivo upon
 CC receiving energy from the Ca(2+)-activated photoprotein aequorin.
 CC Absorbs light maximally at 395 nm and exhibits a smaller
 CC absorbance peak at 470 nm. The fluorescence emission spectrum
 CC peaks at 509 nm with a shoulder at 540 nm.
 CC -1- SUBUNIT: Monomer.
 CC -1- TISSUE SPECIFICITY: Photocytes.
 CC -1- PTM: Contains a covalently attached chromophore, which is composed
 CC of modified amino acid residues. The chromophore is formed upon
 CC cyclization of the residues Ser-dehydrotyr-Gly.
 CC -1- BIOTECHNOLOGY: Has become a useful and ubiquitous tool for making
 CC chimeric proteins of GFP linked to other proteins where it
 CC functions as a fluorescent protein tag. GFP tolerates N- and C-
 CC terminal fusion to a broad variety of proteins. It has been
 CC expressed in bacteria, yeast, slime mold, plants, Drosophila,
 CC zebrafish, and in mammalian cells. As a noninvasive fluorescent
 CC marker in living cells, it allows for a wide range of applications
 CC where it may function as a cell lineage tracer, reporter of gene
 CC expression, or as a measure of protein-protein interactions.
 CC -1- DATABASE: NAME-Protein Spotlight;
 CC NOTE=Issue 11 of June 2001;
 CC WWW="http://www.expasy.org/spotlight/articles/spclt011.html".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/annouce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M62654; AAA27722.1; -;
 DR EMBL; M62653; AAA27721.1; -;
 DR EMBL; L29345; AAA58246.1; -;
 DR EMBL; X96418; CAA65278.1; -;
 DR PIR; J50692; J01514.
 DR PDB; 1B9C; 17-NOV-00.
 DR PDB; 1BFP; 07-JUL-97.
 DR PDB; 1C4F; 14-JUN-00.
 DR PDB; 1EMA; 08-NOV-96.
 DR PDB; 1EMG; 16-JUN-97.
 DR PDB; 1EMH; 20-AUG-97.
 DR PDB; 1EME; 20-AUG-97.
 DR PDB; 1EMF; 20-AUG-97.
 DR PDB; 1EMK; 20-AUG-97.
 DR PDB; 1EML; 20-AUG-97.
 DR PDB; 1EMM; 20-AUG-97.
 DR PDB; 1F03; 17-NOV-00.
 DR PDB; 1F0B; 17-NOV-00.
 DR PDB; 1GFL; 11-JAN-97.
 DR PDB; 1HCU; 15-JAN-97.
 DR PDB; 1HUY; 04-JUL-01.
 DR PDB; 1JBY; 07-JAN-03.
 DR PDB; 1JBE; 07-JAN-03.
 DR PDB; 1KBP; 28-AUG-02.
 DR PDB; 1KYP; 10-APR-02.

DR PDB; 1KXR; 10-APR-02.
 DR PDB; 1KXS; 10-APR-02.
 DR PDB; 1YFP; 28-OCT-98.
 DR PDB; 2EMD; 20-AUG-97.
 DR PDB; 2EMN; 20-AUG-97.
 DR PDB; 2EMO; 20-AUG-97.
 DR InterPro; IPR009017; GFP like.
 DR InterPro; IPR000786; Green_fl_protein.
 DR Pfam; PF01353; GFP; 1.
 DR PRINTS; PR01229; GFPUSCENSANT.
 DR ProDom; PD013756; Green_fl_protein; 1.
 DR Luma; 3D-structure.
 KW CROSSLINK
 FT MOD RES 65 67 5-imidazolinone (Ser-Gly).
 FT MOD RES 66 66 F -> Y.
 FT VARIANT 100 100 T -> S.
 FT VARIANT 108 108 L -> M.
 FT VARIANT 141 141 V -> I.
 FT VARIANT 219 219 S -> G (IN REF. 3).
 FT CONFLICT 2 2 H -> Q (IN REF. 2).
 FT CONFLICT 25 25 Q -> R (IN REF. 3).
 FT CONFLICT 80 80 Q -> P (IN REF. 2).
 FT CONFLICT 157 157 E -> K (IN REF. 2).
 FT CONFLICT 172 172
 FT HELIX 4 8
 FT STRAND 12 22
 FT TURN 23 24
 FT STRAND 25 36
 FT TURN 37 40
 FT STRAND 41 48
 FT TURN 49 50
 FT HELIX 57 60
 FT TURN 61 63
 FT HELIX 69 71
 FT STRAND 73 73
 FT HELIX 76 81
 FT HELIX 83 86
 FT TURN 87 90
 FT STRAND 92 100
 FT TURN 101 102
 FT STRAND 105 115
 FT TURN 116 117
 FT STRAND 118 128
 FT TURN 132 133
 FT STRAND 135 139
 FT TURN 141 141
 FT STRAND 148 155
 FT TURN 156 159
 FT STRAND 160 171
 FT TURN 172 173
 FT STRAND 176 187
 FT STRAND 199 208
 FT TURN 211 212
 FT STRAND 217 227
 SQ SEQUENCE 238 AA; 26886 MW; EA5A6F21FBFB6E05 CRC64;
 Query Match 83.4%; Score 1251; DB 1; Length 238;
 Best Local Similarity 98.3%; Pred. No. 6.2e-94;
 Matches 234; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 2 VSKGEELFTGVVPIIVHLDQVNGHKFSVSGEBSGDTYKGLTLKFLCTTGKLPVWPPTL 61
 DB 1 MSKGEELFTGVVPIIVHLDQVNGHKFSVSGEBSGDTYKGLTLKFLCTTGKLPVWPPTL 60
 QY 62 VTLTYGVQCFSRYPHMKQHPFKSAMPGYQERTIFPKDGNKTRAEVFEFGDTLY 121
 DB 61 VTTFYGVQCFSRYPHMKQHPFKSAMPGYQERTIFPKDGNKTRAEVFEFGDTLY 120
 QY 122 NRTELKGDPEKEDNIIIGHKLEYNNSHNHYIMADKQKGIKYNFKIRHNIEDGSVQLAD 181
 DB 121 NRTELKGDPEKEDNIIIGHKLEYNNSHNHYIMADKQKGIKYNFKIRHNIEDGSVQLAD 180
 QY 182 HYQONTPIGDGPVLIPNNHLSQOSALSKDPNKRBDHMTLEFVTAAGTIDGDELYK 239

DB 181 HYQNTPIDGDPVLLPDPNHYLSTQSALSKDPNEKRDHMLLEFVTAAGITGHMDLYK 238

RESULT 2
DOOR_MOUSE STANDARD; PRT; 461 AA.
AC P00860; Q61997; Q61998;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ornithine decarboxylase (EC 4.1.1.17) (ODC).
GN ODC1 OR ODC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8516193; PubMed=3856848;
RA Kahana C., Nachans D.;
RT "Nucleotide sequence of murine ornithine decarboxylase mRNA."; Proc. Natl. Acad. Sci. U.S.A. 82:1673-1677(1985).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=BAB/c;
RX MEDLINE=88203223; PubMed=3362685;
RA Coffino P., Chen E.L.;
RT "Nucleotide sequence of the mouse ornithine decarboxylase gene."; Nucleic Acids Res. 16:2731-2731(1988).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=85131068; PubMed=2982844;
RA Gupta M., Coffino P.;
RT "Mouse ornithine decarboxylase. Complete amino acid sequence deduced from cDNA."; J. Biol. Chem. 260:2941-2944(1985).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=88227956; PubMed=3372502;
RA Katz A., Kahana C.;
RT "Isolation and characterization of the mouse ornithine decarboxylase gene."; J. Biol. Chem. 263:7604-7609(1988).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=93353527; PubMed=8350350;
RA Johannes G.J., Berger F.G.;
RT "Domains within the mammalian ornithine decarboxylase messenger RNA have evolved independently and episodically."; J. Mol. Evol. 36:555-567(1993).
RN [6]
RP SEQUENCE OF 178-461 FROM N.A.
RX MEDLINE=86120987; PubMed=3456155;
RA Hickok N.J., Seppanen P.J., Kontula K.K., Jaanne P.A.,
RA Bardin C.W., Jaanne O.A.;
RT "Two ornithine decarboxylase mRNA species in mouse kidney arise from size heterogeneity at their 3' termini."; Proc. Natl. Acad. Sci. U.S.A. 83:594-598(1986).
RN [7]
RP PHOSPHORYLATION BY CKII.
RX MEDLINE=91224134; PubMed=2026163;
RA Rosenbergs-Hasson Y., Strumpf D., Kahana C.;
RT "Mouse ornithine decarboxylase is phosphorylated by casein kinase-II at a predominant single location (serine 303)."; Eur. J. Biochem. 197:419-424(1991).
RN [8]
RP ACTIVE SITE, AND PYRIDOXAL-PHOSPHATE BINDING SITE.
RX MEDLINE=92112641; PubMed=1730582;
RA Boulin R., Lu L., Ackermann B., Bey P., Pegs A.E.;
RT "Mechanism of the irreversible inactivation of mouse ornithine decarboxylase by alpha-difluoromethylornithine. Characterization of sequences at the inhibitor and coenzyme binding sites."; J. Biol. Chem. 267:150-158(1992).

RN [9]
RP ACTIVE SITE.
RX MEDLINE=93277871; PubMed=8504104;
RA Tobias K.E., Kahana C.;
RT "Intersubunit location of the active site of mammalian ornithine decarboxylase as determined by hybridization of site-directed mutants."; Biochemistry 32:5842-5847(1993).
RN [10]
RP MUTAGENESIS OF GLY-387.
RX MEDLINE=94062835; PubMed=8243470;
RA Tobias K.E., Mamroud-Kidron E., Kahana C.;
RT "Gly387 of murine ornithine decarboxylase is essential for the formation of stable homodimers."; Eur. J. Biochem. 218:245-250(1993).
RN [11]
RP X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).
RX MEDLINE=99306039; PubMed=10378276;
RA Kern A.D., Oliveira M.A., Coffino P., Hackert M.L.;
RT "Structure of mammalian ornithine decarboxylase at 1.6-A resolution: stereochemical implications of PUF-dependent amino acid decarboxylases."; Structure 7:567-581(1999).
RN [12]
RP CATALYTIC ACTIVITY: L-ornithine = putrescine + CO(2).
CC -1- COFACTOR: Pyridoxal phosphate.
CC -1- PATHWAY: Polyamine biosynthesis; first (rate-limiting) step.
CC -1- SUBUNIT: Homodimer.
CC -1- SIMILARITY: BELONGS TO FAMILY 2 OF ORNITHINE, DAP, AND ARGININE DECARBOXYLASES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; M20617; AAA51638.1; -
DR EMBL; M10624; AAA39845.1; -
DR EMBL; X07392; CAA30301.1; -
DR EMBL; J03733; AAA39849.1; -
DR EMBL; S64539; AAB27809.1; -
DR EMBL; M12330; AAA39846.1; -
DR EMBL; M12331; AAA39848.1; -
DR PIR; A01077; DCM50.
DR PIR; I56477; I56477.
DR PDB; 7ODC; 22-OCT-99.
DR MGD; MGI:97402; Odc.
DR InterPro; IPR001083; Decarboxylase2.
DR InterPro; IPR009006; Racem decarbox_C.
DR Pfam; PF02784; Orn_Arg_dec_N; 1.
DR Pfam; PF02784; Orn_DAP_Arg_dec; 1.
DR PRINTS; PR01179; ODADCBXBLASE.
DR PROSITE; PS00878; ODR_DC_2_1; 1.
DR PROSITE; PS00879; ODR_DC_2_2; 1.
KW Lyase; Decarboxylase; Pyridoxal phosphate; Polyamine biosynthesis; phosphorylation; 3D-structure.
KW BINDING 69 69
FT ACT_SITE 360 360
FT MOD_RES 303 303
FT MUTAGEN 387 387
FT MUTAGEN 387 387
FT CONFLICT 178 178
FT CONFLICT 206 206
FT CONFLICT 350 350
FT STRAND 3 6
FT TURN 7 8
FT STRAND 9 14
FT HELIX 20 28
FT STRAND 40 44
FT HELIX 45 58

PHOSPHORYLATION (BY CK2).
G->A: PARTIAL LOSS OF ACTIVITY.
G->G,D,E,F,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y:
LOS OF ACTIVITY
R->W (IN REF. 6; AAA39846).
D->E (IN REF. 6; AAA39846).
Y->H (IN REF. 6; AAA39848).

```

FT TURN 60 61
FT STRAND 62 67
FT HELIX 68 70
FT HELIX 74 83
FT TURN 84 84
FT STRAND 86 89
FT HELIX 92 100
FT TURN 101 102
FT HELIX 105 107
FT STRAND 108 110
FT HELIX 117 125
FT STRAND 126 127
FT STRAND 130 133
FT HELIX 136 145
FT TURN 147 148
FT STRAND 150 155
FT STRAND 172 172
FT HELIX 174 186
FT TURN 187 188
FT STRAND 190 195
FT TURN 205 205
FT HELIX 206 225
FT TURN 226 226
FT STRAND 231 233
FT HELIX 248 252
FT HELIX 265 267
FT TURN 268 273
FT STRAND 270 273
FT HELIX 277 280
FT HELIX 281 283
FT STRAND 284 296
FT STRAND 313 318
FT TURN 322 326
FT HELIX 327 331
FT STRAND 339 340
FT TURN 346 347
FT STRAND 350 356
FT TURN 362 363
FT STRAND 365 373
FT TURN 377 378
FT STRAND 380 383
FT HELIX 391 393
FT HELIX 397 399
FT STRAND 404 410
FT HELIX 411 417
SQ SEQUENCE 461 AA; 5163 MW; DEF1753FA2ACA0A30 CRC64;

Query Match
Best Local Similarity 87.8%; Score 221.5; DB 1; Length 461;
Matches 43; Conservative 3; Mismatches 2; Indels 1; Gaps 1;

QY 234 MDLYKKL-SHGPPPEVEEDDGLTFMSCAQSSEGMGRHPACASARINV 281
DB 413 MWQMKQIQSHGFPEVEEQDDGLTFMSCAQSSEGMGRHPACASARINV 461

RESULT 3
DCOR MUSPA STANDARD; PRT; 461 AA.
AC P27119;
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ornithine decarboxylase (EC 4.1.1.17) (ODC).
GN ODC1 OR ODC.
OS Mus pahari (Shrew mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10093;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92250571; PubMed=1374399;
RA Johannes G., Berger F.G.;

```

```

RT "Alterations in mRNA translation as a mechanism for the modification
RT of enzyme synthesis during evolution. The ornithine decarboxylase
RT model."
RL J. Biol. Chem. 267:10108-10115 (1992).
CC -1- CATALYTIC ACTIVITY: L-ornithine = putrescine + CO(2).
CC -1- COFACTOR: Pyridoxal phosphate.
CC -1- PATHWAY: Polyamine biosynthesis; first (rate-limiting) step.
CC -1- SUBUNIT: Homodimer.
CC -1- SIMILARITY: BELONGS TO FAMILY 2 OF ORNITHINE, DAP, AND ARGININE
CC DECARBOXYLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M87223; AAA9847.1; -.
DR PIR; I55356; I55356.
DR HSSP; P00860; 7ODC.
DR MGD; MGI:97402; Odc.
DR InterPro; IPR00183; Decarboxylase2.
DR InterPro; IPR009006; Racem decarbox_C.
DR Pfam; PF02784; Orn_Arg_dec_N; 1.
DR Pfam; PF00278; Orn_DAP_Arg_dec; 1.
DR PRINTS; PR01179; ODACRBLXSLR.
DR PROSITE; PS00878; ODR_DC_2_1; 1.
DR PROSITE; PS00879; ODR_DC_2_2; 1.
DR Lyase; Decarboxylase; Pyridoxal phosphate; Polyamine biosynthesis;
KW phosphorylation.
KW BINDING 69 69 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
FT ACT SITE 360 360 BY SIMILARITY.
FT MOD_RES 303 303 PHOSPHORYLATION (BY CK2)
FT (BY SIMILARITY).
SQ SEQUENCE 461 AA; 51203 MW; 61ABCBEC7BFB92 CRC64;

Query Match
Best Local Similarity 29.1%; Score 210.5; DB 1; Length 461;
Matches 78; Conservative 33; Mismatches 70; Indels 87; Gaps 12;

QY 90 PEGYVQ---ERTTFKDDGNYKTRAEYKFGSDTLVNFELKG-----IDPKDGN 136
DB 205 PETFQAVSDARCVF--DMG-----TEVGFs---MYLDDIGGFGPSGSDTYLKEEITLS 253
QY 137 ILGHKLE-YNNNSHNYTMADKKMGKIKVNFKIRHNI-----EDGSYQL 179
DB 254 VINPALDKTFPSDSGVRILAEKRYVASAFTLAVNIATKKTWKEQPSDEDSNBQT 313
QY 180 ADHYQANTPIGDGPVTLDPNHYLSTQSALSKDPNKK-----RDHMY----- 220
DB 314 FMYYVNDGVYSFNCILVDHNVKALLQKRPKPEKYYSSSIWPTCGDLRIYRCMLP 373
QY 221 -----LLE---FTTAAGITLG-----MDLYKKL-SHGPPPEVEED 253
DB 374 EHWVGDWMLFENMGAYTYAAASTFNGFORPNIIYYMSRPMWQMKRIQSHGFPEVEED 433
QY 254 DGLTFMSCAQSSEGMGRHPACASARINV 281
DB 434 DGLTFMSCAQSSEGMGRHPACASARINV 461

RESULT 4
DCOR RAT STANDARD; PRT; 461 AA.
AC P09057;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ornithine decarboxylase (EC 4.1.1.17) (ODC).
GN ODC1 OR ODC.
OS Rattus norvegicus (Rat).

```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Mistar; TISSUE=Liver;
RX MEDLINE=88335556; PubMed=3419906;
RA van Steeg H., van Oostrom C.T.M., van Kranen H.J., van Kreijl C.F.;
RT "Nucleotide sequence of the rat ornithine decarboxylase gene.";
RL Nucleic Acids Res. 16:8173-8174 (1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=86167817; PubMed=3443298;
RA van Kranen H.J., van de Zande L., van Kreijl C.F., Bisschop A.,
RT "Cloning and nucleotide sequence of rat ornithine decarboxylase
RT cDNA.";
RL Gene 60:145-155 (1987).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Fischer; TISSUE=Liver;
RX MEDLINE=89255378; PubMed=2722815;
RA Wen L., Huang J.K., Blackshear P.J.;
RT "Rat ornithine decarboxylase gene. Nucleotide sequence, potential
RT regulatory elements, and comparison to the mouse gene.";
RL J. Biol. Chem. 264:9016-9021 (1989).
CC -1 CATALYTIC ACTIVITY: L-ornithine = putrescine + CO(2).
CC -1 COFACTOR: Pyridoxal phosphate.
CC -1 PATHWAY: Polyamine biosynthesis; first (rate-limiting) step.
CC -1 SUBUNIT: Homodimer.
CC -1 SIMILARITY: BELONGS TO FAMILY 2 OF ORNITHINE, DAP, AND ARGININE
CC DECARBOXYLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, X07944; CAA30765.1; -
DR EMBL, M16982; AAA41737.1; -
DR EMBL, J04792; AAA6286.1; ALT_SEQ.
DR EMBL, J04791; AAA6164.1; -
DR PIR, A27361; DCRTD.
DR HSP, P00860; 70DC.
DR InterPro; IPR00183; Decarboxylase2.
DR InterPro; IPR009006; Racem decarbox_C.
DR Pfam; PF02784; Orn_Arg_dec_N; 1.
DR PRINTS; PR01179; ODAACBXLASE.
DR PROSITE; PS00878; ODR_DC_2_1; 1.
DR PROSITE; PS00879; ODR_DC_2_2; 1.
DR Lyase; Decarboxylase; Pyridoxal phosphate; Polyamine biosynthesis;
KW Phosphorylation.
FT BINDING 69 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
FT ACT_SITE 360 BY SIMILARITY.
FT MOD_RES 303 PHOSPHORYLATION (BY CK2)
FT (BY SIMILARITY).
SQ SEQUENCE 461 AA; 51047 MW; CEF9D268DDDD0AD6 CRC64;
Query Match 13.9%; Score 208.5; DB 1; Length 461;
Best Local Similarity 29.1%; Pred. No. 1.9e-09;
Matches 78; Conservative 33; Mismatches 70; Indels 87; Gaps 12;
QY 90 PEGVQ---ERTFFQDQGNKYRAEYKFGDTLVNRIELKQ-----IDFKEDGN 136
DB 205 PETHVQVSDPARCVF--DMG-----TEVGFs-----MYLLDGGGPGSEDTKLFEEBITs 253
QY 137 ILGHELE-YVNSHNVTIMDKONGIKVNFKIRHNI-----EDGSVOL 179
DB 254 VINPAUKTPPSDSGVATIAFGKRYVASAFTLVANIIAKTWKEQTGSDEDSNBQT 313

```

```

QY 180 ADHYQNTPIGDGVILLPDNHYLSTQSALSKDPNEK-----RDHW----- 220
DB 314 LMTYVNDGVYSGENCILYDAHVAFALIQRPKDEKYSISWGPTCDGLDRIVERCSLP 373
QY 221 -----LLE-----FTVAAGITL-----GMDELYKKL-SHGFPPEVEED 253
DB 374 EMHVGDMMLTENMGATVVAALSTFNGQRPNITYYWSRSMWQMLKQIQSHGFPEVEED 433
QY 254 DGTLPMSCAQESGMDRHPAACASARINV 281
DB 434 VGTLPMSCAQESGMDRHPAACASASINV 461
RESULT 5
DCOR CRIGR STANDARD; PRT; 455 AA.
AC P14019;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ornithine decarboxylase (EC 4.1.1.17) (ODC).
GN ODC1 OR ODC.
OS Cricetus griseus (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetulus.
OX NCBI_TaxID=10029;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9098890; PubMed=2602162;
RA Grens A., Steglich C., Pilz R., Scheffler I.E.;
RT "Nucleotide sequence of the Chinese hamster ornithine decarboxylase
RL Nucleic Acids Res. 17:10497-10497 (1989).
RN [2]
RP SEQUENCE OF 163-455 FROM N.A.
RX MEDLINE=87308329; PubMed=2887574;
RA Strinvaasan P.R., Tonin P.N., Wensing E.J., Lewis W.H.;
RT "The gene for ornithine decarboxylase is co-amplified in hydroxyurea-
RT resistant hamster cells.";
RL J. Biol. Chem. 262:12871-12878 (1987).
CC -1 CATALYTIC ACTIVITY: L-ornithine = putrescine + CO(2).
CC -1 COFACTOR: Pyridoxal phosphate.
CC -1 PATHWAY: Polyamine biosynthesis; first (rate-limiting) step.
CC -1 SUBUNIT: Homodimer.
CC -1 SIMILARITY: BELONGS TO FAMILY 2 OF ORNITHINE, DAP, AND ARGININE
CC DECARBOXYLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, X16910; CAA34784.1; -
DR EMBL, J02813; AAA36963.1; -
DR PIR, S09574; DCHYOC.
DR HSP, P00860; 70DC.
DR InterPro; IPR00183; Decarboxylase2.
DR InterPro; IPR009006; Racem decarbox_C.
DR Pfam; PF02784; Orn_Arg_dec_N; 1.
DR PRINTS; PR01179; ODAACBXLASE.
DR PROSITE; PS00878; ODR_DC_2_1; 1.
DR PROSITE; PS00879; ODR_DC_2_2; 1.
DR Lyase; Decarboxylase; Pyridoxal phosphate; Polyamine biosynthesis;
KW Phosphorylation.
FT BINDING 67 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
FT ACT_SITE 354 BY SIMILARITY.
FT MOD_RES 297 PHOSPHORYLATION (BY CK2)

```

```

FT      CONFLICT      452      452      (BY SIMILARITY)
FT      SEQUENCE      455 AA; 50453 MW; 14162985F18C89F CRC64;
SQ      SEQUENCE      455 AA; 50453 MW; 14162985F18C89F CRC64;

Query Match      13.3%; Score 200; DB 1; Length 455;
Best Local Similarity 28.7%; Pred. No. 9.2e-09;
Matches 76; Conservative 34; Mismatches 71; Indels 84; Gaps 12;

QY      PEGYVQ-----ERTIFPKDGNKYTRAAYVEEGDTLVNRIELKG-----IDFKEDGN 136
DB      PETFYQALSDACV--DMG-----TEVGF--MYLLDIOGGPFGSDTLTKKEELIS 250
QY      ILGHTLE-YNVNSHNVYTMADKXKNGIKYKFKIRINI-----EDGSYQLADH 182
DB      VLNPAIDKFKPPDSGVRVIAEPRGYVASAFTLVNIIAKIKVSKSDDEDESSQTFMY 310
QY      183 YQONTPIGGPVLDPDNHILSTQSALSKDPNEK-----RDHMV----- 220
DB      311 YVNDGYGSENCILYDHAHVKKLPPRPKPKDEKYYSSSTWFTCDGLDIVERCNLPENH 370
QY      221 -----LLE-----FYTAAGITLG-----MDLYEKL-SHGFPPEVERODDGT 256
DB      371 VGDWMLFEMMGAYTYTAAASTFNGFQRPSTIYVMSRPWQMLMQIONHGFPEVEEDVGT 430
QY      257 LPMSCAOSGSMRHPAACASARINV 281
DB      431 LPISCAOSGMDRHPAACASASINV 455

RESULT 6
DCOR HUMAN
ID      DCOR HUMAN      STANDARD;      PRT;      461 AA.
AC      P11926;
DT      01-OCT-1989 (Rel. 12, Created)
DT      01-OCT-1989 (Rel. 12, Last sequence update)
DT      10-OCT-2003 (Rel. 42, Last annotation update)
DE      Ornithine decarboxylase (EC 4.1.1.17) (ODC).
GN      ODC1.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX      NCBI_TaxID:9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=90067851; PubMed=2587220;
RA      van Steeg H., van Oostrom C.T.M., Martens J.W.M., van Kreyl C.F.,
RA      Schepens J., Wieringa B.;
RT      "Nucleotide sequence of the human ornithine decarboxylase gene.";
RL      Nucleic Acids Res. 17:8855-8856(1989).
RN      [2]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=9012632; PubMed=2693021;
RA      Fitzgerald M.C., Flanagan M.A.;
RT      "Characterization and sequence analysis of the human ornithine
RT      decarboxylase gene.";
RL      DNA 8:623-634(1989).
RN      [3]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=87246067; PubMed=3595418;
RA      Hickok N.J., Seppanen P.J., Gunsalus G.L., Jaenne O.A.;
RT      "Complete amino acid sequence of human ornithine decarboxylase
RT      deduced from complementary DNA.";
RL      DNA 6:179-187(1987).
RN      [4]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=91033036; PubMed=2227439;
RA      Hickok N.J., Wahlfors J., Crozat A., Halmekytoe M., Alhonen L.,
RA      Jaenne O.A.;
RT      "Human ornithine decarboxylase-encoding loci: nucleotide sequence of
RT      the expressed gene and characterization of a pseudogene.";
RL      Gene 93:257-263(1990).
RN      [5]
RP      SEQUENCE FROM N.A.

```

```

RX      MEDLINE=92285144; PubMed=1598217;
RA      Moshier J.A., Osborne D.L., Skunca M., Dosescu J., Gilbert J.D.,
RA      Fitzgerald M.C., Polidori G., Wagner R.L., Fritzen Degen S.J.,
RA      Luk G.D., Flanagan M.A.;
RT      "Multiple promoter elements govern expression of the human ornithine
RT      decarboxylase gene in colon carcinoma cells.";
RL      Nucleic Acids Res. 20:2581-2590(1992).
RN      [6]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=90202959; PubMed=2318872;
RA      Moshier J.A., Gilbert J.D., Skunca M., Dosescu J., Almodovar K.M.,
RA      Luk G.D.;
RT      "Isolation and expression of a human ornithine decarboxylase gene.";
RL      J. Biol. Chem. 265:4884-4892(1990).
RN      [7]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=22386257; PubMed=12477932;
RA      Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA      Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA      Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA      Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA      Ditchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,
RA      Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA      Brownstein M.J., Ueda T.B., Tomihata S., Carninci P., Prange C.,
RA      Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA      Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA      Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultky S.W.,
RA      Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA      Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA      Whiting M., Madan A., Young A.C., Shvachenko Y., Boulford G.G.,
RA      Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA      Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA      Butlerfield Y.S.N., Krzywinski M.I., Skalska V., Smallus D.E.,
RA      Scherch A., Schein J.R., Jones S.J.M., Marra M.A.;
RT      "Generation and initial analysis of more than 15,000 full-length
RT      human and mouse cDNA sequences.";
RL      Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN      [8]
RP      SEQUENCE OF 1-22 FROM N.A.
RX      MEDLINE=90199754; PubMed=2317811;
RA      Hsieh J.T., Denning M.F., Heidell S.M., Verna A.K.;
RT      "Expression of human chromosome 2 ornithine decarboxylase gene in
RT      ornithine decarboxylase-deficient Chinese hamster ovary cells.";
RL      Cancer Res. 50:2239-2244(1990).
RN      [9]
RP      X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).
RX      MEDLINE=20090911; PubMed=10623504;
RA      Almud J.O., Oliveira M.A., Kern A.D., Grishin N.V.,
RA      Phillips M.A., Hacker M.L.;
RT      "Crystal structure of human ornithine decarboxylase at 2.1 A
RT      resolution: structural insights to antizyme binding.";
RL      J. Mol. Biol. 295:7-16(2000).
CC      -I- CATALYTIC ACTIVITY: L-ornithine = putrescine + CO(2).
CC      -I- CORFACTOR: Pyridoxal phosphate.
CC      -I- PATHWAY: Polyamine biosynthesis, first (rate-limiting) step.
CC      -I- SUBUNIT: Homodimer.
CC      -I- SIMILARITY: BELONGS TO FAMILY 2 OF ORNITHINE, DAP, AND ARGININE
CC      DECARBOXYLASES.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.ebi.ac.uk/announcements
CC      or send an email to license@ebi.ac.uk).
CC      -----
DR      EMBL; M31061; AAA60563.1; -.
DR      EMBL; M16650; AAA59966.1; -.
DR      EMBL; X16277; CAA34353.1; -.
DR      EMBL; M33764; AAA60564.1; -.
DR      EMBL; X55362; CAA39047.1; -.

```

DR EMBL; M81740; AAA59967.1; -.
 DR EMBL; M34158; AAA59969.1; -.
 DR EMBL; BC025296; AAH25296.1; -.
 DR EMBL; X53271; CA37369.1; -.
 DR PIR; S06900; DCHDO.
 DR PDB; 1D7K; 25-OCT-00.
 DR Genew; HGNC:8109; ODC1.
 DR GK; P11926; -.
 DR MIM; 165640; -.
 DR GO; GO:0004586; F:ornithine decarboxylase activity; TAS.
 DR GO; GO:0006596; P:polyamine biosynthesis; NAS.
 DR InterPro; IPR00183; Decarboxylase.
 DR InterPro; IPR009006; Racem decarbox_C.
 DR Pfam; PF02784; Orn_Arg_dec_N; 1.
 DR Pfam; PF00278; Orn_DAP_Arg_dec; 1.
 DR PRINTS; PR01179; ODADECBLASE.
 DR PROSITE; PS00878; ODR_DC_2_1; 1.
 DR PROSITE; PS00879; ODR_DC_2_2; 1.
 DR Lyase; Decarboxylase; Pyridoxal phosphate; Polyamine biosynthesis;
 KM Phosphorylation; 3D-structure.
 FT BINDING 69 69 PYRIDOXAL PHOSPHATE.
 FT ACT_SITE 360 360 BY SIMILARITY.
 FT MOD_RES 303 303 PHOSPHORYLATION (BY CK2)
 (BY SIMILARITY).
 Q -> E (IN REF. 3).
 FT CONFLICT 415 415
 FT STRAND 11 13
 FT TURN 16 17
 FT HELIX 20 28
 FT TURN 29 30
 FT STRAND 40 44
 FT HELIX 45 58
 FT TURN 60 61
 FT STRAND 62 67
 FT HELIX 68 70
 FT TURN 74 83
 FT HELIX 84 84
 FT STRAND 86 89
 FT HELIX 92 99
 FT TURN 100 102
 FT HELIX 105 107
 FT STRAND 108 110
 FT HELIX 117 125
 FT TURN 126 127
 FT STRAND 130 133
 FT HELIX 136 145
 FT TURN 147 148
 FT STRAND 150 155
 FT STRAND 172 172
 FT HELIX 174 186
 FT TURN 187 188
 FT STRAND 190 195
 FT TURN 200 201
 FT TURN 205 205
 FT HELIX 206 225
 FT TURN 226 226
 FT STRAND 231 233
 FT HELIX 248 262
 FT STRAND 265 267
 FT STRAND 270 273
 FT HELIX 277 280
 FT HELIX 281 283
 FT STRAND 284 296
 FT STRAND 313 318
 FT TURN 322 326
 FT HELIX 327 331
 FT STRAND 339 340
 FT TURN 346 347
 FT STRAND 350 356
 FT TURN 362 363
 FT STRAND 365 373
 FT TURN 377 378
 FT STRAND 380 383
 FT HELIX 391 393

FT HELIX 397 399
 FT STRAND 404 409
 FT HELIX 411 420
 SQ SEQUENCE 461 AA; 51148 MW; 8CCB8CE80823C5 CRC64;
 Query Match 10.3%; Score 154.5; DB 1; Length 461;
 Best Local Similarity 27.2%; Pred. No. 4.4e-05;
 Matches 72; Conservative 29; Mismatches 83; Indels 81; Gaps 12;
 QY 90 PEEVQ-----ERIFRDOONNYTRAAYKE-----GDTLVNRLEKIGDFEDNGNITLG 139
 DB 205 PEEVQAISDARCVF--DWG-----AEVGFMYLIDIGGFPSSBDVK-LKFEELIGVIN 256
 QY 140 HKLE-YVNSHNYVIMADKXNGIKVNFKIRHNI-----EDSSVQADH 182
 DB 257 PALDKYFPPSSGVRITIAEPERYVAAFTLVNIIAKKIYKQTSDDDESSSQETMY 316
 QY 183 YQONTPIGDGPVLLPNHVLSTQSALSKEPNEK-----RDHNV----- 220
 DB 317 YVNDGVYGFNCILYDHAHVKPLQKRPKPEDEKYSSSIWPTCDGLDRIVERCDLPEMH 376
 QY 221 ----LIE---FVTAAGITLG---MDELYKLS-----HGPPPYEODDGT 256
 DB 377 VGDMLFENNCAATVAASFTNGFORPTIYVMSGPAMQMOQPNPDPPEVEDPAST 436
 QY 257 LPMSCAQESGMDRHPAACAASATNV 281
 DB 437 LPVSCAWESGMRHRAACASATNV 461
 RESULT 7
 ID DCOR BOVIN STANDARD; PRT; 461 AA.
 AC P2717;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Ornithine decarboxylase (EC 4.1.1.17) (ODC).
 GN ODC1 OR ODC.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
 OC Bovidae; Bovinae; Bos.
 ON NCBI_TaxID=9913;
 RX MEDLINE=9293216; Pubmed=7774801;
 RA Yao J., Zadworny D., Kuhnlein U., Hayes J.F.;
 RT "Molecular cloning of a bovine ornithine decarboxylase cDNA and its
 use in the detection of restriction fragment length polymorphisms in
 Holsteins";
 RL Genome 38:325-331(1995).
 CC -1- CATALYTIC ACTIVITY: L-ornithine = putrescine + CO(2).
 CC -1- COFACTOR: Pyridoxal phosphate.
 CC -1- PATHWAY: Polyamine biosynthesis; first (rate-limiting) step.
 CC -1- SUBUNIT: Homodimer.
 CC -1- SIMILARITY: BELONGS TO FAMILY 2 OF ORNITHINE, DAP, AND ARGININE
 DECARBOXYLASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M92441; AAA92339.1; -.
 DR EMBL; U36394; AAA79849.1; -.
 DR EMBL; U18531; AAA6696.1; -.
 DR HSSP; P11926; 1D7K.
 DR InterPro; IPR00183; Decarboxylase2.

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; X56316; CA93760.1; -
 CC PIR; A43563; A43563.
 CC HSSP; P00860; 70DC.
 CC InterPro; IPR000183; Decarboxylase2.
 CC InterPro; IPR000906; Racem decarbox_C.
 CC Pfam; PF02784; Orn_Arg_dec_N; 1.
 CC Pfam; PF00278; Orn_DAP_Arg_dec; 1.
 CC PRINTS; PR01179; ODADCRXLASE.
 CC PROSITE; PS00878; ODR_DC_2_1; 1.
 CC PROSITE; PS00879; ODR_DC_2_2; 1.
 CC Lysase; Decarboxylase; Pyridoxal phosphate; Polyamine biosynthesis;
 CC Phosphorylation.
 CC BINDING 69 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 CC ACT_SITE 361 BY SIMILARITY.
 CC MOD_RES 303 PHOSPHORYLATION (BY CK2)
 CC (BY SIMILARITY).
 CC SEQUENCE 460 AA; 50830 MW; CCGF87FD165877EE CRC64;
 Query Match 6.9%; Score 103; DB 1; Length 460;
 Best Local Similarity 23.8%; Pred. No. 0.63;
 Matches 64; Conservative 34; Mismatches 81; Indels 90; Gaps 14;
 QY 90 PEGYVQ---ERTIFFKDDGNYKRAEYKFGDITVRIELKG-----IDFKENGN 136
 Db 205 PQTVAQVSDARCVF--DMG-----AELGFN---MHLLDGGGPGSGEDVYKLFEEITIS 253
 QY 137 ILHGKLE-YNYSNHYVIMADKKNIGKYNF-----KIRHNIEDGSYQLAD----- 181
 Db 254 VINPALDKYFPADSGVKILAEGRYVASSFTLANIILAKVMNVEGSGSDDEDAANDK 313
 QY 182 ---HYQNTPIGDGEVILPNNHYISTGSAKSDPENK-----RDHNV----- 220
 Db 314 TLMVYVDGVGVSGFNCILFDHAHYVPLTKKPKDEKYSISSIWPGTCDGLDRIYRPEL 373
 QY 221 ---LLE-----FVTAAGITLG---MDLYKKIS-----HGFPPEVDEQ 252
 Db 374 PELQGVDMLEFENMGAYVVAALASTFNGCRPTLYVNSRPHQMMDIKHGIILPEV--P 431
 QY 253 DDGTLPMSCAQSQSGMDRHPAACAASARINV 281
 Db 432 DLSALHVSQAQSGEMELAPVCTAASINV 460
 RESULT 10
 COPA_HUMAN STANDARD; PRT; 1224 AA.
 ID COPA_HUMAN
 AC P53621;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Coatomer alpha subunit (Alpha-coat protein) (Alpha-COP) (HEPCOP)
 DE (HEP-COP) [Contains: Xenin (Xenopsin-related peptide); Proxenin].
 GN COPA.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96194806; PubMed=8647451;
 RA Chow V.T.K., Quek H.H.;
 RT "HEP-COP, a novel human gene whose product is highly homologous to
 RL the alpha-subunit of the yeast coatomer protein complex.";
 RL Gene 169:223-227(1996).
 RN [2]
 RP SEQUENCE OF 1-25 (XENIN).
 RC TISSUE=Gastric mucosa;

RX MEDLINE=93054515; PubMed=1429581;
 RA Feurle G.E., Hamscher G., Kusiek R., Meyer H.E., Metzger J.W.;
 RT "Identification of xenin, a xenopsin-related peptide, in the human
 RL gastric mucosa and its effect on exocrine pancreatic secretion.";
 RL J. Biol. Chem. 267:22305-22309(1992).
 RN [3]
 RP PROCESSING OF COPA TO PRODUCE XENIN.
 RX MEDLINE=98032493; PubMed=9365789;
 RA Chow V.T., Quek H.H.;
 RT "Alpha coat protein COPA (HEP-COP): presence of an Alu repeat in CDNA
 RL and identity of the amino terminus to xenin.";
 RN Ann. Hum. Genet. 61:369-373(1997).
 RN [4]
 RP REVIEW ON XENIN.
 RX MEDLINE=98192336; PubMed=9533652;
 RA Feurle G.E.;
 RT "Xenin -- a review.";
 RL Peptides 19:609-615(1998).
 CC -1- FUNCTION: The coatomer is a cytosolic protein complex that binds
 CC to dylysine motifs and reversibly associates with Golgi non-
 CC clathrin-coated vesicles, which further mediate biosynthetic
 CC protein transport from the ER, via the Golgi up to the trans Golgi
 CC network. Coatomer complex is required for budding from Golgi
 CC membranes, and is essential for the retrograde Golgi-to-ER
 CC transport of dylysine-tagged proteins. In mammals, the coatomer
 CC can only be recruited by membranes associated to ADP-ribosylation
 CC factors (ARFs), which are small GTP-binding proteins; the complex
 CC also influences the Golgi structural integrity, as well as the
 CC processing, activity, and endocytic recycling of LDL receptors (By
 CC similarity).
 CC -1- FUNCTION: XENIN STIMULATES EXOCRINE PANCREATIC SECRETION. IT
 CC INHIBITS PENTAGASTRIN-STIMULATED SECRETION OF ACID, TO INDUCE
 CC EXOCRINE PANCREATIC SECRETION AND TO AFFECT SMALL AND LARGE
 CC INTESTINAL MOTILITY. IN THE GUT, XENIN INTERACTS WITH THE
 CC NEUTROPHILIN RECEPTOR.
 CC -1- SUBUNIT: Oligomeric complex that consists of at least the alpha,
 CC beta, beta', gamma, delta, epsilon and zeta subunits.
 CC -1- SUBCELLULAR LOCATION: THE COATOMER IS CYTOPLASMIC OR POLYMERIZED
 CC ON THE CYTOPLASMIC SIDE OF THE GOLGI, AS WELL AS ON THE
 CC VESICLES/BUDS ORIGINATING FROM IT (BY SIMILARITY). XENIN IS
 CC SECRETED.
 CC -1- TISSUE SPECIFICITY: UNIFORMLY EXPRESSED IN A WIDE RANGE OF ADULT
 CC AND FETAL TISSUES. XENIN IS FOUND IN GASTRIC, DODENAL AND JECUNAL
 CC MUCOSA. CIRCULATES IN THE BLOOD. SEEMS TO BE CONFINED TO SPECIFIC
 CC ENDOCRINE CELLS.
 CC -1- DEVELOPMENTAL STAGE: XENIN IS RELEASED INTO THE CIRCULATION AFTER
 CC A MEAL.
 CC -1- SIMILARITY: Contains 6 WD repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U24105; AAB70879.1; -
 CC PIR; J04668; ERHDAH.
 CC Genew; H8NC; 2230; COPA.
 CC MIM; 601924; -
 CC GO; GO:0005803; C:secretory vesicle; TAS.
 CC GO; GO:0005215; F:transporter activity; TAS.
 CC GO; GO:0006888; P:ER to Golgi transport; TAS.
 CC InterPro; IPR006692; Coatomer WDAD.
 CC InterPro; IPR008941; TPR-like.
 CC InterPro; IPR001880; WD40.
 CC Pfam; PF04053; Coatomer_WDAD; 1.
 CC Pfam; PF00400; WD40; 6.
 CC PRINTS; PR00320; GPROTEINRPT.
 CC PRODOM; PD000018; WD40; 5.
 CC SMART; SMO0320; WD40; 7.
 CC PROSITE; PS00678; WD_REPEATS_1; 1.

DR PROSITE; PS50082; WD_REPEATS_2; 6.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Transports; Protein transport; Golgi stack; Endoplasmic reticulum;
KW Membrane; Repeat; WD repeat; Hormone.
FT PEPTIDE 1 25 XENIN.
FT REPEAT 1 35 PRORENIN.
FT REPEAT 7 37 WD 1.
FT REPEAT 49 79 WD 2.
FT REPEAT 91 121 WD 3.
FT REPEAT 133 163 WD 4.
FT REPEAT 203 233 WD 5.
FT REPEAT 247 277 WD 6.
SQ SEQUENCE 1224 AA; 138331 MW; 5A8E884D58F155D CRC64;

Query Match
Best Local Similarity 26.6%; Score 99.5; DB 1; Length 1224;
Matches 47; Conservative 28; Mismatches 68; Indels 37; Gaps 8;

QY 91 EGYVQERTFFPDGNGYKTRAEVKEGDTLVNRIR-LKGDIFKEDGNILGHKLEYNNSH 149
DB 632 KGYPEVALHFKVDE---KTRFSIALBEGCNIEALBAKALDDKNCWEMKGEVALIQGNHQ 688
QY 150 NVYIMADKQKNGIKYVNF-----KIRHNIEDGSVQ--LADHYQONTPIGSGVLLPD 198
DB 689 IYEMCYQRTKNDKXSFYLITGNIEKLRKMKKIAIRKDSGHIQNALYIGD----- 741
QY 199 NHYLSTQSALSKDNEKRDHNVLLFEYTAGITLGMDELKYLKSHGPPVEVBQDDGTLF 258
DB 742 ---VSEKRYILKNGGKS-----LAVLTAA--THGLDEBAESLKETPDPEK-----TIP 786

RESULT 11
COPA_BOVIN
AC Q27954; STANDARD; PRT; 1224 AA.
ID 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Coatomer alpha subunit (Alpha-coat protein) (Alpha-COP) (HEPCOP)
DR (HEP-COP) [Contains: Xenin (Xenopsin-related peptide); Prorenin].
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
NX NCBI_TaxID=9913;
RN SEQUENCE FROM N.A.
RA Wieland F.I.;
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: The coatomer is a cytosolic protein complex that binds to dilysine motifs and reversibly associates with Golgi non-clathrin-coated vesicles, which further mediate biosynthetic protein transport from the ER, via the Golgi up to the trans Golgi network. Coatomer complex is required for budding from Golgi membranes, and is essential for the retrograde Golgi-to-ER transport of dilysine-tagged proteins. In mammals, the coatomer can only be recruited by membranes associated to ADP-ribosylation factors (ARFs), which are small GTP-binding proteins; the complex also influences the Golgi structural integrity, as well as the processing, activity, and endocytic recycling of LDL receptors (By similarity).
CC -!- FUNCTION: XENIN STIMULATES EXOCRINE PANCREATIC SECRETION. IT INHIBITS PENTAGASTRIN-STIMULATED SECRETION OF ACID, TO INDUCE EXOCRINE PANCREATIC SECRETION AND TO AFFECT SMALL AND LARGE INTESTINAL MOTILITY. IN THE GUT, XENIN INTERACTS WITH THE NEUROGENSIN RECEPTOR (BY SIMILARITY).
CC -!- SUBUNIT: Oligomeric complex that consists of at least the alpha, beta, beta', gamma, delta, epsilon and zeta subunits.
CC -!- SUBCELLULAR LOCATION: THE COATOMER IS CYTOSOLIC OR POLYMERIZED ON THE CYTOSOLIC SIDE OF THE GOLGI, AS WELL AS ON THE VESICLES/BUDS ORIGINATING FROM IT (BY SIMILARITY). XENIN IS SECRETED (BY SIMILARITY).

CC -!- SIMILARITY: Contains 6 WD repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb.ch/announce/> or send an email to license@isb-sib.ch).
CC EMBL; X96768; CAA65543.1; -
DR InterPro; IPR006692; Coatomer_WDAD.
DR InterPro; IPR008941; TPR-1like.
DR InterPro; IPR001680; WD40.
DR Pfam; PF04053; Coatomer_WDAD; 1.
DR Pfam; PF00400; WD40; 6.
DR PRINTS; PR00320; GPROTEINRPT.
DR PRODOM; PD000018; WD40; 5.
DR SMART; SM00320; WD40; 7.
DR PROSITE; PS00678; WD_REPEATS_1; 1.
DR PROSITE; PS50082; WD_REPEATS_2; 6.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Transports; Protein transport; Golgi stack; Endoplasmic reticulum;
KW Membrane; Repeat; WD repeat; Hormone.
FT PEPTIDE 1 25 XENIN.
FT REPEAT 1 35 PRORENIN.
FT REPEAT 7 37 WD 1.
FT REPEAT 49 79 WD 2.
FT REPEAT 91 121 WD 3.
FT REPEAT 133 163 WD 4.
FT REPEAT 203 233 WD 5.
FT REPEAT 247 277 WD 6.
SQ SEQUENCE 1224 AA; 138358 MW; C9350BF2AC00683D CRC64;

Query Match
Best Local Similarity 25.6%; Score 96.5; DB 1; Length 1224;
Matches 46; Conservative 29; Mismatches 68; Indels 37; Gaps 8;

QY 91 EGYVQERTFFPDGNGYKTRAEVKEGDTLVNRIR-LKGDIFKEDGNILGHKLEYNNSH 149
DB 632 KGYPEVALHFKVDE---KTRFSIALBEGCNIEALBAKALDDKNCWEMKGEVALIQGNHQ 688
QY 150 NVYIMADKQKNGIKYVNF-----KIRHNIEDGSVQ--LADHYQONTPIGSGVLLPD 198
DB 689 IYEMCYQRTKNDKXSFYLITGNIEKLRKMKKIAIRKDSGHIQNALYIGD----- 741
QY 199 NHYLSTQSALSKDNEKRDHNVLLFEYTAGITLGMDELKYLKSHGPPVEVBQDDGTLF 258
DB 742 ---VSEKRYILKNGGKS-----LAVLTAA--THGLDEBAESLKETPDPEK-----TIP 786

RESULT 12
NECL_RAT
ID P28840; STANDARD; PRT; 752 AA.
AC 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Neuroendocrine convertase 1 precursor (EC 3.4.21.93) (PCL)
DE (Prohormone convertase 1) (Proprotein convertase 1).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NX NCBI_TaxID=10116;
RN SEQUENCE FROM N.A.
RX MEDLINE=92168040; PubMed=1791845;
RA Bloomquist B.T., Eipper B.A., Mains R.E.;
RT "Prohormone-converting enzymes: regulation and evaluation of function using antisense RNA."
RL Mol. Endocrinol. 5:2014-2024 (1991).


```

RP SEQUENCE FROM N.A.
RX MEDLINE=92063860; PubMed=1954888;
RA Hakes D.J., Birch N.P., Mezey A., Dixon J.E.;
RT "Isolation of two complementary deoxyribonucleic acid clones from a
RT rat insulinoma cell line based on similarities to Kex2 and furin
RT sequences and the specific localization of each transcript to
RT endocrine and neuroendocrine tissues in rats.";
RL Endocrinology 129:3053-3063(1991).
CC -1- FUNCTION: Involved in the processing of hormone and other protein
CC precursors at sites comprised of pairs of basic amino acid
CC residues. Substrates include POMC, renin, enkephalin, dynorphin,
CC somatostatin and insulin.
CC -1- CATALYTIC ACTIVITY: Release of protein hormones, neuropeptides and
CC renin from their precursors, generally by cleavage of -Lys-Arg-|-
CC bonds.
CC -1- COFACTOR: Calcium.
CC -1- SUBCELLULAR LOCATION: Localized in the secretion granules.
CC -1- SIMILARITY: Belongs to peptidase family S8. Furin subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M75705; AAA40945.1; -
DR EMBL: M83745; AAA41476.1; -
DR PIR: A41556; KXRTCL.
DR HSSP: Q45670; IDBI.
DR MEROPS: S08.072; -
DR InterPro: IPR002029; Peptidase S8.
DR InterPro: IPR002884; Peptidase_S8B.
DR InterPro: IPR009020; Protease_inhib.
DR Pfam: PF01483; P_proprotein; 1.
DR Pfam: PF00082; Peptidase_S8; 1.
DR PRINTS: PR00723; SUBTILISIN.
DR ProDom: PD000717; P_domain; 1.
DR PROSITE: PS00136; SUBTILASE ASP; 1.
DR PROSITE: PS00137; SUBTILASE_HIS; 1.
DR PROSITE: PS00138; SUBTILASE_SER; 1.
DR Hydrolase; Serine protease; Glycoprotein; Zymogen; Calcium; Signal.
FT SIGNAL 1 27
FT PROPEP 28 110
FT CHAIN 111 752
FT DOMAIN 122 410
FT DOMAIN 739 751
FT ACT_SITE 167 167
FT ACT_SITE 208 382
FT ACT_SITE 382 382
FT CARBOHYD 173 173
FT CARBOHYD 401 401
FT CARBOHYD 645 645
FT CONFLICT 366 366
FT CONFLICT 514 514
SQ SEQUENCE 752 AA; 84120 MW; F630AD830A076DED CRC64;

Query Match 6.2%; Score 92.5; DB 1; Length 752;
Best Local Similarity 24.4%; Pred. No. 8.1;
Matches 60; Conservative 29; Mismatches 80; Indels 77; Gaps 13;

18 ELGDVNGHKK-----FSVSGEGEDATYGLTKLFICTTGKLP-----VPMPTLVTLTYG 68
Db 540 ERLTSPRGFMQPMFMSHTMGENV--GTWTLKATLDMSGMONGRIVNKKLL-----HG 593
QY 69 VQCESRPDMKQKQDFKSMPEGYVOERITTFKDDGNVYTRAEVKEGDTLVNRIELKG 128
Db 594 T---SSQPEHMKQ-----PRYTL-----SYNTVQDRRGVKKMNVVVEKP 631
QY 129 IDFKEDNIIIGHLEEVYNSHNYIYADKQKNGIKVFKIRHNIEDSSVOLADHYOONT 188
Db 632 TQNSLNGNLIVPK---NSSSSVVEDRDRDQVQCAPSKAMLR-----LIGSAFSKNIP 680

```

```

QY 189 IGDGPIVLPDNNHYLSTQSAISKDPENKRDHMYLFEVTAAGITAGMDLYLKKLSHGFPPE 248
Db 681 -----SKQS-----SKIPSKLS-----VPIEGLYEALFKLNK-----PSQ 711
QY 249 VERQDD 254
Db 712 LQESD 717

RESULT 13
ID YD48 METUA STANDARD; PRT; 336 AA.
AC OS8743;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein MJ1348.
GN MJ1348.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2461 / ATCC 43067;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.L.,
RA Overbeek R., Kirkness E.F., Weisskopf K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Ulfreback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurt M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii.";
RL Science 273:1058-1073(1996).
RN [2]
RP DISCUSSION ON SEQUENCE.
RX MEDLINE=98362752; PubMed=9697413;
RA Kyriades N.C., Woese C.R.;
RT "Tetratric-peptide-repeat proteins in the archaeon Methanococcus
RT jannaschii.";
RL Trends Biochem. Sci. 23:245-247(1998).
CC -1- CAUTION: According to Ref.2, this sequences contains TPR repeats.
CC These are not detected using our methodology.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U67574; AAB99360.1; -
DR PIR: C64468; C64468.
DR TIGR: MJ1348; -
DR InterPro: IPR008941; TPR-like.
DR Hypothetical protein; Complete proteome.
SQ SEQUENCE 336 AA; 40443 MW; 41BF2F28CDB036FE CRC64;

Query Match 6.1%; Score 92; DB 1; Length 336;
Best Local Similarity 24.9%; Pred. No. 3.3;
Matches 66; Conservative 37; Mismatches 80; Indels 82; Gaps 17;

QY 20 DGDVNGHKSVSGE-----GGDATYGLT-----TLKFICTTGKLPVFW-----PTL 61
Db 91 DGDY-----VNLSEGLSTIASIFAKIGKLDITKNFNSGGE-----WYNVYIKDANSEDL 142
QY 62 VTLLTYGV-----QCFSRYPDMKQ-----HDFPKSMPEGYVOERITTFKDDG 105

```

Db 143 KSVLTFFDSYEERKEILNREP-HLRKLFEDNTYNENSDFFEDFM-----MFFIGAG 192
 QY 106 NKXTAAYKFE-----GDTLVNRLTELKIDPEKEDNTIL--GHKLEYVYNHANYI 153
 Db 193 NKRRLEAYEEFKKIKSCQISNEIYNEL-IRFPD-KMSMDLALAHLLKENYECCLYIV 250
 QY 154 YADK-----QKNGIKVNFKIRHNIEDSGVQADHYOONTPIGDGPAVLDP-DNHY 201
 Db 251 MLFFKEFPLDEPNEIEKKNNILIVLAVLKNQNVKKEEMLNRLNEIYKEIKREPLPMY 310
 QY 202 ISTQSALSKPD-PNEKRDHNVLLIEFV 225
 Db 311 ---KDAHNDLNLNEILDVYVLAKEFI 332

RESULT 14

TRMA CAMJE STANDARD; PRT; 357 AA.
 ID Q9P92;

AC 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE tRNA (Uracil-5)-methyltransferase (EC 2.1.1.35) (tRNA (M-5-US4) -methyltransferase) (RUMT).
 GN TRMA OR C00831C.

OS Campylobacter jejuni.
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
 OC Campylobacteraceae; Campylobacter.

OX NCBI_Taxid=197;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=NCCTC 11168;

RA MEDLINE=20150912; PubMed=1068204; Kelley J.M., Churcher C.,
 RA Parthill J., Wren B.W., Mungall K., Davies R.M., Feltwell T., Holroyd S.,
 RA Basham D., Chillingworth T., Davies R.M., Feltwell T., Holroyd S.,
 RA Jagels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
 RA Quail M.A., Rajandream M.A., Rutherford K.K., van Vliet A.H.M.,
 RA Whitehead S., Barrett B.G.;
 RA "The genome sequence of the food-borne pathogen Campylobacter jejuni
 RT reveals hypervariable sequences.";
 RL Nature 403:665-668 (2000).

CC -!- FUNCTION: Catalyzes the formation of 5-methyl-uridine at position
 CC 54 (M-5-US4) in all tRNA (by similarity).
 CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + tRNA = S-adenosyl-L-
 CC homocysteine + tRNA containing thymine.
 CC -!- SIMILARITY: Belongs to the RNA M5U methyltransferase family. Trna
 CC subfamily.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: AL139076; CAB7096.1; -.
 DR HAMAP: G81355; G81355.

DR HAMAP: G81355; G81355.
 DR InterPro: IPR000051; SAM bind.

DR InterPro: IPR001566; Trna.
 DR PROSITE: PS01230; TRMA_1; 1.

DR PROSITE: PS01231; TRMA_2; FALSE_NEG.
 DR TRANSFERASE: Methyltransferase; tRNA processing; Complete proteome.

DR TRANSFERASE: Methyltransferase; tRNA processing; Complete proteome.
 FT DOMAIN 207 213

FT ACT SITE 315 315 BY SIMILARITY.
 FT SEQUENCE 357 AA; 42276 MW; CEC5328347CEB97 CRC64;

Query Match 6.1%; Score 91; DB 1; Length 357;
 Best Local Similarity 19.5%; P-Value 4.3;
 Matches 53; Conservative 43; Mismatches 84; Indels 92; Gaps 11;

QY 80 KOHDFEFSAMEBEGYVQERTIFFKDDGNYKTRAEYKFE--EGTLYV-----121
 Db 14 EKHSFIKKYFEFFETPKFKASDKKHYRRAELSFHENDTLFYAMFDKSKKKYIIIEY 73
 QY 122 -----NRIELK--GIDPEKEDNTILGHKLEYVYN-----147
 Db 74 LDFADEKICAMFPLLELYLQDNKLEKELFGVEELTQKEISTLLVHKNIEDIKSNLEN 133
 QY 148 -SHNVYIMADQKNGKIGVNEKIRH-----NIEDSGVQADHYOONTPIGDGPAVLDPNHY 201
 Db 134 LSNLTHTLTLARSGKGLIFETENVRQTLNTIQDKI-----FYEPN-----NDCE 178
 QY 202 ISTQSALSKPD-----RDHNVLLIEFVTAAG-ITLGMDELKYLKSHGFPPEV 249
 Db 179 IQPNTAL-----NEWMITWVCEILNTQKRMDELLEYCGVGNFTLALAPFFKI---LATIEI 231
 QY 250 EEDDGTIPMSCAQBSGKDRHPACASARINV 281
 Db 232 SKSNINPALKNCELNNTNTHPARLSSEBEL 263

RESULT 15

CP51 CANCA STANDARD; PRT; 533 AA.
 ID CP51 CANCA

AC P50859; Q02312; (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Cytochrome P450 51 (EC 1.14.13.70) (CYPLI) (P450-LI1) (Sterol 14-
 DE alpha-demethylase) (lanosterol 14-alpha demethylase) (P450-14DM).
 GN ERG1 OR CYP51.

OS Candida glabrata (Yeast) (Torulopsis glabrata).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; mitosporic Saccharomycetales; Candida.

OX NCBI_Taxid=5478;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=2001-L5;

RA MEDLINE=96161286; PubMed=8593007;
 RA Geber A., Hitchcock C.A., Swartz J.E., Pullen F.S., Marsden K.E.,
 RA Kwon-Chung K.J., Bennett J.E.;
 RA "Deletion of the Candida glabrata ERG3 and ERG1 genes: effect on cell
 RT viability, cell growth, sterol composition, and antifungal
 RT susceptibility.";
 RL Antimicrob. Agents Chemother. 39:2708-2717 (1995).

CC -!- FUNCTION: Catalyzes C14-demethylation of lanosterol which is
 CC critical for ergosterol biosynthesis. It transforms lanosterol
 CC into 4,4'-dimethyl cholesterol-8,14,24-triene-3-beta-ol (By
 CC similarity).
 CC -!- CATALYTIC ACTIVITY: Obusufolol + 3 O(2) + 3 NADPH = 4-alpha-
 CC methyl-5-alpha-ergosta-8,14,24-trien-3-beta-ol + formate + 3
 CC NADP(+) + 3 H(2)O.

CC -!- PATHWAY: Ergosterol biosynthesis.
 CC -!- SIMILARITY: Belongs to the cytochrome P450 family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

```

CC -----
DR EMBL; L40389; AAB02329.1; -.
DR EMBL; S75389; AAB32679.1; -.
DR InterPro; IPR01128; Cytochrome_P450.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
KW Electron transport; Oxidoreductase; Monooxygenase; Membrane; Heme;
KW Sterol biosynthesis; NADP.
FT METAL 472 472 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
FT CONFLICT 64 64 I -> M (IN REF. 2).
FT CONFLICT 473 473 I -> T (IN REF. 2).
SQ SEQUENCE 533 AA; 61305 MW; A0506C17507B6E7 CRC64;

Query Match 6.1%; Score 91; DB 1; Length 533;
Best Local Similarity 20.3%; Pred. No. 7;
Matches 54; Conservative 44; Mismatches 106; Indels 62; Gaps 11;

QY 25 GHKESVS---GEGGDATYGGKLTIKFTCTTGKLPVPMFTLVTTLYGVQCFSRYPDH--M 79
Db 109 GHEFIFFNAKLADVSAEAAYSHL-----TTPVFGKVIYDCPNHRLM 149
QY 80 KOHDFEKSAM--PEGVY-----OERTIEFKDDGNVYKTRAEVKEGDTLVNRIELKGDIF 131
Db 150 EQKFFVKGALYKEAFVRYVPLIAEIIYKFRNSKPFKINENSGIVDWSQPEM--TIF 207
QY 132 KEDGNILGHKLEVYNSHNYIMADKQNGIKVNFKIRHNIEDGSVQLADHYQQTPIGD 191
Db 208 TASRSLGKKEWRDLDTDFAYLSDLDKGFPIINF--VEPNPLAEHYRRDHAQAIS--- 263
QY 192 GPVLLPDNNHYISTQSALSKDPNEKEDHWLLEFVTAAGITLGMDELYKGLSHGPPPEVEE 251
Db 264 -----GTWMSLKERERKND-----IONRDL---IDELMKNSTYKDGTYKMTD 302
QY 252 QDDGTLPMSCAQSAGMDRHPAACA 277
Db 303 QETANLLIGVIMG---QHTSAATSA 325

```

Search completed: May 14, 2004, 08:25:46
 Job time : 20 secs

Db 61 VTTFSYGVQCSRRYDHNKRDHFKSAMPEGYVQERTIFFPKDGNKYTRAIVEKGGDTLV 120
 QY 122 NRIELKIGIDFKEDGNILGHKLEYNNSHNHYIMADKONGIKVNFKHNIHEDGSVOLAD 181
 Db 121 NRIELKIGIDFKEDGNILGHKLEYNNSHNHYIMADKONGIKVNFKHNIHEDGSVOLAD 180
 QY 182 HYQONTPIGDSGVLLPDNNHYLSTQSALSKDPNEKRDHMLLEFVTAAGITGMDELYK 239
 Db 181 HYQONTPIGDSGVLLPDNNHYLSTQSALSKDPNEKRDHMLLEFVTAAGITGMDELYK 238

RESULT 2

Q93125 PRELIMINARY; PRT; 238 AA.
 AC Q93125;
 DT 01-FEB-1997 (Tremblrel. 02, Created)
 DT 01-FEB-1997 (Tremblrel. 02, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE Green fluorescent protein mutant 3.
 GN GFP.
 OS Aequorea victoria (Jellyfish).
 OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydrozoa; Leptomedusae;
 OC Aequoridae; Aequorea.
 OX NCBI_TaxID=6100;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96305137; PubMed=8707053;
 RA Cormack B.P., Valdivia R.H., Falkow S.;
 RT "FACS-optimized mutants of the green fluorescent protein (GFP).";
 RL Gene 173:33-38 (1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Cormack B.P., Berttram G., Egerton M., Gow N.A.R., Falkow S.,
 RA Brown A.J.P.;
 RT "Yeast Enhanced Green Fluorescent Protein (yEGFP): a reporter of gene
 expression in Candida albicans.";
 RL Microbiology 0:0-0 (1996).
 DR EMBL; U73901; AAB18957.1; -.
 DR HSSP; P42312; 1BFP.
 DR GO; GO:0006091; P:energy pathways; IEA.
 DR InterPro; IPR009017; GFP_like.
 DR InterPro; IPR00786; Green_fl_protein.
 DR Pfam; PF01353; GFP; 1.
 DR PRINTS; PRO1229; GFPLORESCENT.
 DR ProDom; PD013756; Green_fl_protein; 1.
 SQ SEQUENCE 238 AA; 26840 MW; A28622809A9DEA6q CRC64;

Query Match 83.0%; Score 1245; DB 5; Length 238;
 Best Local Similarity 97.9%; Pred. No. 5.4e-95;

Matches 233; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 VSKGEELFTGVVPIVLVDGVDNKHKFSVSGEGDATYGLTKLFICTTGKLPVPMPTL 61
 Db 1 MSKGEELFTGVVPIVLVDGVDNKHKFSVSGEGDATYGLTKLFICTTGKLPVPMPTL 60
 QY 62 VTTLYGVQCSRRYDHNKRDHFKSAMPEGYVQERTIFFPKDGNKYTRAIVEKGGDTLV 121
 Db 61 VTTFSYGVQCSRRYDHNKRDHFKSAMPEGYVQERTIFFPKDGNKYTRAIVEKGGDTLV 120
 QY 122 NRIELKIGIDFKEDGNILGHKLEYNNSHNHYIMADKONGIKVNFKHNIHEDGSVOLAD 181
 Db 121 NRIELKIGIDFKEDGNILGHKLEYNNSHNHYIMADKONGIKVNFKHNIHEDGSVOLAD 180
 QY 182 HYQONTPIGDSGVLLPDNNHYLSTQSALSKDPNEKRDHMLLEFVTAAGITGMDELYK 239
 Db 181 HYQONTPIGDSGVLLPDNNHYLSTQSALSKDPNEKRDHMLLEFVTAAGITGMDELYK 238

RESULT 3

Q8GHE4 PRELIMINARY; PRT; 238 AA.
 AC Q8GHE4;
 DT 01-MAR-2003 (Tremblrel. 23, Created)

DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE Green fluorescence protein.
 GN 375GFP.
 OS Azomonas agilis.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Azomonas.
 OX NCBI_TaxID=116849;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Koranyi P., Berenyi M., Burg K.;
 RT "Occurrence of green fluorescence protein in diazotrophic bacteria
 RT Azomonas and Azotobacter";
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF324405; AAN86137.1; -.
 DR GO; GO:0006091; P:energy pathways; IEA.
 DR InterPro; IPR009017; GFP_like.
 DR InterPro; IPR00786; Green_fl_protein.
 DR Pfam; PF01353; GFP; 1.
 DR PRINTS; PRO1229; GFPLORESCENT.
 DR ProDom; PD013756; Green_fl_protein; 1.
 SQ SEQUENCE 238 AA; 26902 MW; 15FE9B9C5B4F6B89 CRC64;

Query Match 82.8%; Score 1242; DB 2; Length 238;
 Best Local Similarity 97.5%; Pred. No. 9.5e-95;

Matches 232; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 VSKGEELFTGVVPIVLVDGVDNKHKFSVSGEGDATYGLTKLFICTTGKLPVPMPTL 61
 Db 1 MSKGEELFTGVVPIVLVDGVDNKHKFSVSGEGDATYGLTKLFICTTGKLPVPMPTL 60
 QY 62 VTTLYGVQCSRRYDHNKRDHFKSAMPEGYVQERTIFFPKDGNKYTRAIVEKGGDTLV 121
 Db 61 VTTFSYGVQCSRRYDHNKRDHFKSAMPEGYVQERTIFFPKDGNKYTRAIVEKGGDTLV 120
 QY 122 NRIELKIGIDFKEDGNILGHKLEYNNSHNHYIMADKONGIKVNFKHNIHEDGSVOLAD 181
 Db 121 NRIELKIGIDFKEDGNILGHKLEYNNSHNHYIMADKONGIKVNFKHNIHEDGSVOLAD 180
 QY 182 HYQONTPIGDSGVLLPDNNHYLSTQSALSKDPNEKRDHMLLEFVTAAGITGMDELYK 239
 Db 181 HYQONTPIGDSGVLLPDNNHYLSTQSALSKDPNEKRDHMLLEFVTAAGITGMDELYK 238

RESULT 4

Q8GHE3 PRELIMINARY; PRT; 238 AA.
 AC Q8GHE3;
 DT 01-MAR-2003 (Tremblrel. 23, Created)
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE Green fluorescence protein.
 GN 85GFP.
 OS Azotobacter vinelandii.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Azotobacter.
 OX NCBI_TaxID=354;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Koranyi P., Berenyi M., Burg K.;
 RT "Occurrence of green fluorescence protein in diazotrophic bacteria
 RT Azomonas and Azotobacter";
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF324406; AAN86138.1; -.
 DR GO; GO:0006091; P:energy pathways; IEA.
 DR InterPro; IPR009017; GFP_like.
 DR InterPro; IPR00786; Green_fl_protein.
 DR Pfam; PF01353; GFP; 1.
 DR PRINTS; PRO1229; GFPLORESCENT.
 DR ProDom; PD013756; Green_fl_protein; 1.
 SQ SEQUENCE 238 AA; 26887 MW; B0E1616BD2AF6188 CRC64;

Query Match 82.7%; Score 1241; DB 2; Length 238;

Best Local Similarity 97.5%; Pred. No. 1,1e-94;
Matches 232; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

```

QY 2 VSKSEELFTGVVPLIVELDGVNGHKFSVSGEGDATTGKLTLLKFTCTTGKLPVPPPTL 61
D 1 MSKSEELFTGVVPLIVELDGVNGHKFSVSGEGDATTGKLTLLKFTCTTGKLPVPPPTL 60
QY 62 VTTLTYGVQCFSRYPDMKQDFFKSAMPEGYOERTIFFKDDGNKYTRAEVKEGGTLY 121
D 61 VTTSYGVQCFSRYPDMKQDFFKSAMPEGYOERTIFFKDDGNKYTRAEVKEGGTLY 120
QY 122 NRILKIGIDPEKEDNIIIGHKLEYNNSHNVIYIMADKQNGIKVNFRIHNTEDGSVOLAD 181
D 121 NRILKIGIDPEKEDNIIIGHKLEYNNSHNVIYIMADKQNGIKVNFRIHNTEDGSVOLAD 180
QY 182 HYQONTPIGDGPVLLPDNHYLSTQSALSKDNEKRDMVLLFEVTAAGITLGMDELYK 239
D 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDNEKRDMVLLFEVTAAGITLGMDELYK 238

```

RESULT 5

```

ID 017105 PRELIMINARY; PRT; 238 AA.
AC 017105;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Green fluorescent protein (Fragment).
GN GFP.
OS Aequorea victoria (Jellyfish).
OC Eukaryota; Metazoa; Chordata; Hydrozoa; Hydroida; Leptomedusae;
OC Aequoreidae; Aequorea.
OX NCBI_TaxID=6100;
RN [1]
RE SEQUENCE FROM N.A.
RA Watkins J.N., Campbell A.K.;
RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; X83959; CA58789.1; -.
DR PIR; J50692; J01514.
DR HSSP; P42212; 1BFP.
DR GO; GO:0006091; P:energy pathways; IEA.
DR InterPro; IPR009017; GFP_like.
DR InterPro; IPR00786; Green_fl_protein.
DR Pfam; PF01353; GFP; 1.
DR PRINTS; PRO1229; GFPDORSCENT.
DR ProDom; PD013756; Green_fl_protein; 1.
FT NON_TER 238
SQ SEQUENCE 238 AA; 26950 MM; 26E2BE450E748B44 CRC64;

```

Query Match 80.6%; Score 1209; DB 5; Length 238;
Best Local Similarity 94.1%; Pred. No. 5,1e-92;
Matches 224; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

```

QY 2 VSKSEELFTGVVPLIVELDGVNGHKFSVSGEGDATTGKLTLLKFTCTTGKLPVPPPTL 61
D 1 MSKSEELFTGVVPLIVELDGVNGHKFSVSGEGDATTGKLTLLKFTCTTGKLPVPPPTL 60
QY 62 VTTLTYGVQCFSRYPDMKQDFFKSAMPEGYOERTIFFKDDGNKYTRAEVKEGGTLY 121
D 61 VTTSYGVQCFSRYPDMKQDFFKSAMPEGYOERTIFFKDDGNKYTRAEVKEGGTLY 120
QY 122 NRILKIGIDPEKEDNIIIGHKLEYNNSHNVIYIMADKQNGIKVNFRIHNTEDGSVOLAD 181
D 121 NRILKIGIDPEKEDNIIIGHKLEYNNSHNVIYIMADKQNGIKVNFRIHNTEDGSVOLAD 180
QY 182 HYQONTPIGDGPVLLPDNHYLSTQSALSKDNEKRDMVLLFEVTAAGITLGMDELYK 239
D 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDNEKRDMVLLFEVTAAGITLGMDELYK 238

```

RESULT 6
017106 PRELIMINARY; PRT; 238 AA.

```

AC 017106;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Green fluorescent protein (Fragment).
GN GFP.
OS Aequorea victoria (Jellyfish).
OC Eukaryota; Metazoa; Chordata; Hydrozoa; Hydroida; Leptomedusae;
OC Aequoreidae; Aequorea.
OX NCBI_TaxID=6100;
RN [1]
RE SEQUENCE FROM N.A.
RA Watkins J.N., Campbell A.K.;
RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; X83960; CA58790.1; -.
DR PIR; J50692; J01514.
DR HSSP; P42212; 1BFP.
DR GO; GO:0006091; P:energy pathways; IEA.
DR InterPro; IPR009017; GFP_like.
DR InterPro; IPR00786; Green_fl_protein.
DR Pfam; PF01353; GFP; 1.
DR PRINTS; PRO1229; GFPDORSCENT.
DR ProDom; PD013756; Green_fl_protein; 1.
FT NON_TER 238
SQ SEQUENCE 238 AA; 26867 MM; BD4648262DBEABD4 CRC64;

```

Query Match 79.6%; Score 1194; DB 5; Length 238;
Best Local Similarity 93.3%; Pred. No. 8,9e-91;
Matches 222; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

```

QY 2 VSKSEELFTGVVPLIVELDGVNGHKFSVSGEGDATTGKLTLLKFTCTTGKLPVPPPTL 61
D 1 MSKSEELFTGVVPLIVELDGVNGHKFSVSGEGDATTGKLTLLKFTCTTGKLPVPPPTL 60
QY 62 VTTLTYGVQCFSRYPDMKQDFFKSAMPEGYOERTIFFKDDGNKYTRAEVKEGGTLY 121
D 61 VTTSYGVQCFSRYPDMKQDFFKSAMPEGYOERTIFFKDDGNKYTRAEVKEGGTLY 120
QY 122 NRILKIGIDPEKEDNIIIGHKLEYNNSHNVIYIMADKQNGIKVNFRIHNTEDGSVOLAD 181
D 121 NRILKIGIDPEKEDNIIIGHKLEYNNSHNVIYIMADKQNGIKVNFRIHNTEDGSVOLAD 180
QY 182 HYQONTPIGDGPVLLPDNHYLSTQSALSKDNEKRDMVLLFEVTAAGITLGMDELYK 239
D 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDNEKRDMVLLFEVTAAGITLGMDELYK 238

```

RESULT 7

```

ID 08WTC6 PRELIMINARY; PRT; 238 AA.
AC 08WTC6;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Green fluorescent protein.
GN GFP.
OS Aequorea macrodactyla.
OC Eukaryota; Metazoa; Chordata; Hydrozoa; Hydroida; Leptomedusae;
OC Aequoreidae; Aequorea.
OX NCBI_TaxID=147615;
RN [1]
RE SEQUENCE FROM N.A.
RA STRAIN=GFPxm19uv;
RA Luo W.X., Zhang J., Yang H.T., Xie X.Y., Li S.W., Guan B.Q., Chen M.,
  Li S.J., Xia N.S.;
RT "Colorful mutants of green fluorescent protein from Aequorea
  macrodactyla.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF435431; AL433916.1; -.
DR GO; GO:0006091; P:energy pathways; IEA.
DR InterPro; IPR009017; GFP_like.
DR InterPro; IPR00786; Green_fl_protein.
DR Pfam; PF01353; GFP; 1.

```

DR PRINTS: PR01229; GFPUSCENST.
DR Pfam: PD013756; Green fl protein; 1.
SQ SEQUENCE 238 AA; 27015 MW; 688FD75B8926903 CRC64;

Query Match 72.6%; Score 1089; DB 5; Length 238;
Best Local Similarity 82.4%; Pred. No. 4.3e-82;
Matches 196; Conservative 21; Mismatches 21; Indels 0; Gaps 0;

QY 2 VSKGEELFTGVVPIVLELDGVNKHKFSVSGEGSDATYGLTKLFTCTTGKLPVPMPTL 61
DB 1 MSKGEELFTGVVPIVLELDGVNKHKFSVSGEGSDATYGLTKLFTCTTGKLPVPMPTL 60
QY 62 VTTLTGVGQCFRSPYEDHMKQDFFKSAMPEGYOERTIFFKDDGNVKTAAVKEGSDTLV 121
DB 61 VTTLTGVGQCFRSPYEDHMKQDFFKSAMPEGYOERTIFFKDDGNVKTAAVKEGSDTLV 120
QY 122 NRIELKIDFKEDGNILGHKLEYNNSHNYIMADKQNGIKVNFKIRHNIEDSGVQLAD 181
DB 121 NRIELKIDFKEDGNILGHKLEYNNSHNYIMADKQNGIKVNFKIRHNIEDSGVQLAD 180
QY 182 HYQNTPIGDGPVLLPNNHYLSTQSALSKDPNEKRDHNVLLFEVTAAGITLGMDELK 239
DB 181 HYQNTPIGDGPVLLPNNHYLSTQSALSKDPNEKRDHNVLLFEVTAAGITLGMDELK 238

RESULT 8

Q8WP95 PRELIMINARY; PRT; 238 AA.

AC Q8WP95;
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Green fluorescent protein.
GN GFPKM.
OS Aequorea macrodactyla.
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydrozoa; Leptomedusae;
OC Aequoreidae; Aequorea.
OX NCBI_TaxID=147615;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-GFPKM, and GFPdnam;
RA Luo W.X., Zhang J., Yang H.J., Xie X.Y., Li S.W., Qian Y.X., Pang S.Q.,
RA Li S.J., Xia N.S.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY013824; AAK02062.1; -
DR EMBL: AY013821; AAK02059.1; -
DR GO: GO:0006091; P:energy pathways; IEA.
DR InterPro: IPR009017; GFP like.
DR Pfam: PF01353; GFP 1.
DR PRINTS: PR01229; GFPUSCENST.
DR Pfam: PD013756; Green fl protein; 1.
SQ SEQUENCE 238 AA; 27049 MW; 8185D0B5E529012B CRC64;

Query Match 72.3%; Score 1085; DB 5; Length 238;
Best Local Similarity 81.9%; Pred. No. 9.2e-82;
Matches 195; Conservative 21; Mismatches 22; Indels 0; Gaps 0;

QY 2 VSKGEELFTGVVPIVLELDGVNKHKFSVSGEGSDATYGLTKLFTCTTGKLPVPMPTL 61
DB 1 MSKGEELFTGVVPIVLELDGVNKHKFSVSGEGSDATYGLTKLFTCTTGKLPVPMPTL 60
QY 62 VTTLTGVGQCFRSPYEDHMKQDFFKSAMPEGYOERTIFFKDDGNVKTAAVKEGSDTLV 121
DB 61 VTTLTGVGQCFRSPYEDHMKQDFFKSAMPEGYOERTIFFKDDGNVKTAAVKEGSDTLV 120
QY 122 NRIELKIDFKEDGNILGHKLEYNNSHNYIMADKQNGIKVNFKIRHNIEDSGVQLAD 181
DB 121 NRIELKIDFKEDGNILGHKLEYNNSHNYIMADKQNGIKVNFKIRHNIEDSGVQLAD 180
QY 182 HYQNTPIGDGPVLLPNNHYLSTQSALSKDPNEKRDHNVLLFEVTAAGITLGMDELK 239
DB 181 HYQNTPIGDGPVLLPNNHYLSTQSALSKDPNEKRDHNVLLFEVTAAGITLGMDELK 238

RESULT 9

Q8WTC4 PRELIMINARY; PRT; 238 AA.

AC Q8WTC4;
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Green fluorescent protein.
GN GFP.
OS Aequorea macrodactyla.
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydrozoa; Leptomedusae;
OC Aequoreidae; Aequorea.
OX NCBI_TaxID=147615;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Shg24;
RA Luo W.X., Zhang J., Yang H.J., Xie X.Y., Li S.W., Qian B.Q., Chen M.,
RA Li S.J., Xia N.S.;
RL "Colorful mutants of green fluorescent protein from Aequorea
macrodactyla.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF435433; AAL33918.1; -
DR GO: GO:0006091; P:energy pathways; IEA.
DR InterPro: IPR009017; GFP like.
DR Pfam: PF01353; GFP 1.
DR PRINTS: PR01229; GFPUSCENST.
DR Pfam: PD013756; Green fl protein; 1.
SQ SEQUENCE 238 AA; 26956 MW; 75521EAF0CEBA73A CRC64;

Query Match 72.1%; Score 1082; DB 5; Length 238;
Best Local Similarity 83.2%; Pred. No. 1.6e-81;
Matches 198; Conservative 17; Mismatches 23; Indels 0; Gaps 0;

QY 2 VSKGEELFTGVVPIVLELDGVNKHKFSVSGEGSDATYGLTKLFTCTTGKLPVPMPTL 61
DB 1 MSKGEELFTGVVPIVLELDGVNKHKFSVSGEGSDATYGLTKLFTCTTGKLPVPMPTL 60
QY 62 VTTLTGVGQCFRSPYEDHMKQDFFKSAMPEGYOERTIFFKDDGNVKTAAVKEGSDTLV 121
DB 61 VTTLTGVGQCFRSPYEDHMKQDFFKSAMPEGYOERTIFFKDDGNVKTAAVKEGSDTLV 120
QY 122 NRIELKIDFKEDGNILGHKLEYNNSHNYIMADKQNGIKVNFKIRHNIEDSGVQLAD 181
DB 121 NRIELKIDFKEDGNILGHKLEYNNSHNYIMADKQNGIKVNFKIRHNIEDSGVQLAD 180
QY 182 HYQNTPIGDGPVLLPNNHYLSTQSALSKDPNEKRDHNVLLFEVTAAGITLGMDELK 239
DB 181 HYQNTPIGDGPVLLPNNHYLSTQSALSKDPNEKRDHNVLLFEVTAAGITLGMDELK 238

RESULT 10

Q8WTD0 PRELIMINARY; PRT; 238 AA.

AC Q8WTD0;
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Green fluorescent protein.
GN GFP.
OS Aequorea macrodactyla.
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydrozoa; Leptomedusae;
OC Aequoreidae; Aequorea.
OX NCBI_TaxID=147615;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GFPxm161;
RA Luo W.X., Zhang J., Yang H.J., Xie X.Y., Li S.W., Qian B.Q., Chen M.,
RA Li S.J., Xia N.S.;
RL "Colorful mutants of green fluorescent protein from Aequorea
macrodactyla.";

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-GFPxmi19uv;
 RA Luo W.X., Zhang J., Yang H.J., Xie X.Y., Li S.W., Guan B.Q., Chen M.,
 RT "Colorful mutants of green fluorescent protein from Aequorea
 macrodactyla."
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF435430; AAL33915.1; -
 DR GO; GO:0006091; P:energy pathways; IEA.
 DR InterPro; IPR009017; GFP like.
 DR InterPro; IPR000786; Green_fl_protein.
 DR Pfam; PF01353; GFP; 1.
 DR PRINTS; PR01229; GFPDORSCENT.
 DR Prodom; PD013756; Green_fl_protein; 1.
 SQ SEQUENCE 238 AA; 27002 MW; BDSBA2982264C018 CRC64;

Query Match 71.7%; Score 1076; DB 5; Length 238;
 Best Local Similarity 81.5%; Pred. No. 5.1e-81;
 Matches 194; Conservative 21; Mismatches 23; Indels 0; Gaps 0;

QY 2 VSKGEELFTGVVPIVLVDGVDNGHKSFSVSGEGDATTGKLTLLKFTCTTGKLPVMPPTL 61
 DB 1 MSKGELFTGIVPVLLIEDGDVHGKFSVRGSGEGDADYGLKIFICTTGKLPVMPPTL 60
 QY 62 VTTLTGVOCFSRYPDHMKQDHPFKSAMPEGVQERTIFFKDDGNKTRAEVKEGDTLV 121
 DB 61 VTTLSYGLICFARYDEHMKMNDFFKSAMPEGYIQERTIFFQDDGKXKTFGEVKEGDTLV 120
 QY 122 NRLEKIGDFEKDNLGKLEYNNSHNYIMADKQNGIKVNFKIRNIEDSGVOLAD 181
 DB 121 NRLEKGMDFEKDNLGKLEYNNSHNYIMPDANGLKVNFKIRNIEGGVOLAD 180
 QY 182 HYQONTPIGDGVLLPDNHYLSTQSALSKDPNEKRDHVLLEFVYTAGITLGMDELK 239
 DB 181 HYQNVPLIGDGVLLPINHYLSQGLAISKDRNETHVFLFFSACGHTHGMDELK 238

RESULT 14
 Q8WTC5 PRELIMINARY; PRT; 238 AA.
 ID Q8WTC5;
 AC Q8WTC5;
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Orange fluorescent protein.
 GN GFP.
 OS Aequorea macrodactyla.
 OC Eukaryota; Metazoa; Chnidaria; Hydrozoa; Hydrozoa; Leptomedusae;
 OC Aequoreidae; Aequorea.
 NC NCB1_TaxID=147615;
 RX [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=OFPxm;
 RA Luo W.X., Zhang J., Yang H.J., Xie X.Y., Li S.W., Guan B.Q., Chen M.,
 RT "Colorful mutants of green fluorescent protein from Aequorea
 macrodactyla."
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF435432; AAL33917.1; -
 DR GO; GO:0006091; P:energy pathways; IEA.
 DR InterPro; IPR009017; GFP like.
 DR InterPro; IPR000786; Green_fl_protein.
 DR Pfam; PF01353; GFP; 1.
 DR PRINTS; PR01229; GFPDORSCENT.
 DR Prodom; PD013756; Green_fl_protein; 1.
 SQ SEQUENCE 238 AA; 27018 MW; 75521BA5534E573A CRC64;

Query Match 71.7%; Score 1075; DB 5; Length 238;
 Best Local Similarity 82.8%; Pred. No. 6.2e-81;
 Matches 197; Conservative 17; Mismatches 24; Indels 0; Gaps 0;
 QY 2 VSKGEELFTGVVPIVLVDGVDNGHKSFSVSGEGDATTGKLTLLKFTCTTGKLPVMPPTL 61

DB 1 MSKGELFTGIVPVLLIEDGDVHGKFSVRGSGEGDADYGLKIFICTTGKLPVMPPTL 60
 QY 62 VTTLTGVOCFSRYPDHMKQDHPFKSAMPEGVQERTIFFKDDGNKTRAEVKEGDTLV 121
 DB 61 VTTLSYGLICFARYDEHMKMNDFFKSAMPEGYIQERTIFFQDDGKXKTFGEVKEGDTLV 120
 QY 122 NRLEKIGDFEKDNLGKLEYNNSHNYIMADKQNGIKVNFKIRNIEDSGVOLAD 181
 DB 121 NRLEKGMDFEKDNLGKLEYNNSHNYIMPDANGLKVNFKIRNIEGGVOLAD 180
 QY 182 HYQONTPIGDGVLLPDNHYLSTQSALSKDPNEKRDHVLLEFVYTAGITLGMDELK 239
 DB 181 HYQNVPLIGDGVLLPINHYLSQGLAISKDRNETHVFLFFSACGHTHGMDELK 238

RESULT 15
 Q95UA7 PRELIMINARY; PRT; 225 AA.
 ID Q95UA7;
 AC Q95UA7;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Cyan fluorescent protein (Fragment).
 OS Montastraea cavernosa (great star coral).
 OC Eukaryota; Metazoa; Chnidaria; Anthozoa; Zoantharia; Scleractinia;
 OC Faviina; Faviidae; Montastraea.
 NC NCB1_TaxID=63558;
 RX [1]
 RN \$SEQUENCE FROM N.A.
 RA Falkowski P.G., Sun Y.;
 RT "Montastraea cavernosa fluorescent protein."
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY056460; AAL17905.1; -
 DR GO; GO:0006091; P:energy pathways; IEA.
 DR InterPro; IPR009017; GFP like.
 DR InterPro; IPR000786; Green_fl_protein.
 DR Pfam; PF01353; GFP; 1.
 DR PRINTS; PR01229; GFPDORSCENT.
 DR Prodom; PD013756; Green_fl_protein; 1.
 FT NON TER 225
 SQ SEQUENCE 225 AA; 25775 MW; 52DE2F716D083524 CRC64;

Query Match 16.8%; Score 251.5; DB 5; Length 225;
 Best Local Similarity 31.1%; Pred. No. 7.8e-13;
 Matches 64; Conservative 43; Mismatches 82; Indels 17; Gaps 7;

QY 12 VPIVLVDGVDNGHKSFSVSGEGDATTGKLTLLKFTCTTGKLPVMPPTLVTTLTYGQ 70
 DB 7 VKIKLRMDGIVNGHKEFKITGBGSKPFGHTTILTKKKEGRLPFAVDILTTFAYQGNR 66
 QY 71 GFSRYPDHMKQDHPFKSAMPEGVQERTIFFKDDGNKTRAEVKEGDTLVNRLEKIGD 130
 DB 67 VFTKYPKQIP--DYRKQSFPEGSWSRSMTEFDQGVCTVTSIDIKLEGSCFFYEIRFYGN 124
 QY 131 FKEDNITLGHK-LEYNNSHNYIMADKQNGIKVNFKIRNIEDSGVOLADHYQONTPI 189
 DB 125 FSSGDPVQKQKTLKMEPSTENNTV-----RDGVLLGDVSRITLLRGD---KHKCNRRS 175
 QY 190 GDGP---VLLPDNHYLSTQ-SALSKD 211
 DB 176 TYGAKKGVLLPEYHFVDHRIETLSHD 201

Search completed: May 14, 2004, 08:26:43
 Job time : 47 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 14, 2004, 08:23:59 ; Search time 22 Seconds

(Without alignments)
659,404 Million cell updates/sec

Title: US-09-931-232-1
Perfect score: 1500
Sequence: 1 MWSKGEFLFGVPIVLELD.....AQSGMDRHPACASARINV 281

Scoring table: BLOSUM62
Gapop 10.0 , Gapect 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/2/iaa/5A COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCBUS COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1500	100.0	281	US-09-062-102-1	Sequence 1, Appl1
2	1500	100.0	281	US-09-364-946-1	Sequence 1, Appl1
3	1289.5	86.0	805	US-09-513-783A-178	Sequence 178, App
4	1280.5	85.4	1171	US-09-417-197-131	Sequence 131, App
5	1279	85.3	997	US-09-417-197-121	Sequence 121, App
6	1278	85.2	607	US-09-417-197-47	Sequence 47, Appl
7	1276	85.1	544	US-09-417-197-115	Sequence 115, Appl
8	1276	85.1	631	US-09-417-197-39	Sequence 39, Appl
9	1276	85.1	843	US-09-417-197-117	Sequence 117, App
10	1275.5	85.0	933	US-09-417-197-135	Sequence 135, App
11	1275.5	85.0	941	US-09-513-783A-172	Sequence 172, App
12	1275	85.0	727	US-09-417-197-139	Sequence 139, App
13	1275	85.0	797	US-09-417-197-143	Sequence 143, App
14	1274	84.9	239	US-09-172-063-3	Sequence 3, Appl1
15	1274	84.9	239	US-09-513-783A-46	Sequence 46, Appl1
16	1274	84.9	239	US-09-316-919-4	Sequence 4, Appl1
17	1274	84.9	239	US-09-602-641-3	Sequence 2, Appl1
18	1274	84.9	239	US-09-920-922-2	Sequence 2, Appl1
19	1274	84.9	294	US-09-513-783A-2	Sequence 2, Appl1
20	1274	84.9	323	US-09-172-063-21	Sequence 21, Appl
21	1274	84.9	323	US-09-602-641-21	Sequence 21, Appl
22	1274	84.9	364	US-09-085-305-6	Sequence 6, Appl1
23	1274	84.9	379	US-09-417-197-129	Sequence 129, App
24	1274	84.9	432	US-09-800-170-48	Sequence 48, Appl
25	1274	84.9	459	US-09-417-197-127	Sequence 127, App
26	1274	84.9	459	US-09-513-783A-170	Sequence 170, App
27	1274	84.9	544	US-09-417-197-113	Sequence 113, App

28	1274	84.9	604	US-09-417-197-59	Sequence 59, Appl
29	1274	84.9	605	US-09-417-197-41	Sequence 41, Appl
30	1274	84.9	606	US-09-417-197-65	Sequence 65, Appl
31	1274	84.9	630	US-09-417-197-63	Sequence 63, Appl
32	1274	84.9	633	US-09-417-197-45	Sequence 45, Appl
33	1274	84.9	635	US-09-417-197-125	Sequence 125, Appl
34	1274	84.9	642	US-08-818-253-6	Sequence 2, Appl1
35	1274	84.9	642	US-08-818-253-2	Sequence 2, Appl1
36	1274	84.9	642	US-08-818-252-6	Sequence 6, Appl1
37	1274	84.9	652	US-08-818-252-4	Sequence 4, Appl1
38	1274	84.9	652	US-08-818-253-4	Sequence 4, Appl1
39	1274	84.9	718	US-09-417-197-75	Sequence 75, Appl
40	1274	84.9	719	US-09-417-197-51	Sequence 51, Appl
41	1274	84.9	726	US-09-417-197-71	Sequence 71, Appl
42	1274	84.9	726	US-09-513-783A-176	Sequence 176, App
43	1274	84.9	793	US-09-417-197-141	Sequence 141, App
44	1274	84.9	797	US-09-417-197-141	Sequence 141, App
45	1274	84.9	798	US-09-417-197-77	Sequence 77, Appl

ALIGNMENTS

RESULT 1
US-09-062-102-1
; Sequence 1, Application US/09062102
; Patent No. 6130313
; GENERAL INFORMATION:
; APPLICANT: Kain, Steve
; APPLICANT: Li, Xiangliang
; TITLE OF INVENTION: Rapidly Degrading GPP-Fusion Proteins and Methods
; FILE REFERENCE: D6100
; CURRENT APPLICATION NUMBER: US/09/062, 102
; CURRENT FILING DATE: 1998-04-17
; EARLIER APPLICATION NUMBER: US 60/060, 855
; EARLIER FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 3
; SEQ ID NO 1
; LENGTH: 281
; TYPE: PRT
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: Sequence of the EGFP-MODC422-461 fusion protein.
; Patent No. 6130313
US-09-062-102-1

Query Match 100.0%; Score 1500; DB 3; Length 281;
Best Local Similarity 100.0%; Pred. No. 1.2e-150;
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MWSKGEFLFGVPIVLELDGVNKHKFSVSGEGEDATYKLTLEKICTTGKLPVWPPT 60
DB 1 MWSKGEFLFGVPIVLELDGVNKHKFSVSGEGEDATYKLTLEKICTTGKLPVWPPT 60
QY 61 LVTTITVGVQCFRSYDPDHMKOHDFKSNAMPEGVQRTIFFKDDGNYKTRAYPEEGDTL 120
DB 61 LVTTITVGVQCFRSYDPDHMKOHDFKSNAMPEGVQRTIFFKDDGNYKTRAYPEEGDTL 120
QY 121 VNIIEKKGIDFKEDGNITGKLENNYNSHNYVIMAKOKNGIKVNRKIRINIEDGSVQLA 180
DB 121 VNIIEKKGIDFKEDGNITGKLENNYNSHNYVIMAKOKNGIKVNRKIRINIEDGSVQLA 180
QY 181 DHVQONTPIGDGVLLPDNNHLYSTQSLAKDPNEKRDHWLLEFVTAAGITLGMDELKXK 240
DB 181 DHVQONTPIGDGVLLPDNNHLYSTQSLAKDPNEKRDHWLLEFVTAAGITLGMDELKXK 240
QY 241 LSHGFPPEVEEDDGTIPMSCAQESGMDRHPACASARINV 281
DB 241 LSHGFPPEVEEDDGTIPMSCAQESGMDRHPACASARINV 281

RESULT 2

US-09-364-946-1
 ; Sequence 1, Application US/09364946
 ; Patent No. 6306600
 ; GENERAL INFORMATION:
 ; APPLICANT: Kain, Steve
 ; APPLICANT: Li, Xiangqiang
 ; TITLE OF INVENTION: Rapidly Degrading GFP-Fusion Proteins and Methods
 ; TITLE OF INVENTION: of Use
 ; FILE REFERENCE: D6100CIP/D2
 ; CURRENT APPLICATION NUMBER: US/09/364,946
 ; EARLIER FILING DATE: 1999-07-30
 ; EARLIER FILING DATE: 1998-11-13
 ; NUMBER OF SEQ ID NOS: 14
 ; SEQ ID NO 1
 ; LENGTH: 281
 ; TYPE: PRT
 ; ORGANISM: artificial sequence
 ; FEATURE:
 ; OTHER INFORMATION: Sequence of the EGFP-MODC422-461 fusion protein.
 ; Patent No. 6306600
 US-09-364-946-1

Query Match 100.0%; Score 1500; DB 4; Length 281;
 Best Local Similarity 100.0%; Pred. No. 1.2e-150;
 Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Cy 1 MWSKGEELFTGVVPIVLELDGVDNGHKFSVSGEGBDATYGLTLTKFICTTGKLPVPMPT 60
    |||
Db 1 MWSKGEELFTGVVPIVLELDGVDNGHKFSVSGEGBDATYGLTLTKFICTTGKLPVPMPT 60

Cy 61 LVTTLYGVOCFSRYPDHMKQHDFFKSAMPEGVVOERTIFFKDDGNVKTAEVKEGDTL 120
    |||
Db 61 LVTTLYGVOCFSRYPDHMKQHDFFKSAMPEGVVOERTIFFKDDGNVKTAEVKEGDTL 120

Cy 121 VNRLEKIDPKEDGNILGHKLEYNNSHNVYIMADKQKGIKVNFKIRHNIEDGSVOLA 180
    |||
Db 121 VNRLEKIDPKEDGNILGHKLEYNNSHNVYIMADKQKGIKVNFKIRHNIEDGSVOLA 180

Cy 181 DHYQONTPIGDGPVLPDNNHYLSTOSALSKDPNEKRDMVLEFVTAAGITLGMDELYKK 240
    |||
Db 181 DHYQONTPIGDGPVLPDNNHYLSTOSALSKDPNEKRDMVLEFVTAAGITLGMDELYKK 240

Cy 241 LSHGFPPEVEBODDGLTPMSCAQSQSGMDRHPAACAASARINV 281
    |||
Db 241 LSHGFPPEVEBODDGLTPMSCAQSQSGMDRHPAACAASARINV 281
  
```

RESULT 3
 US-09-513-783A-178
 ; Sequence 178, Application US/09513783A
 ; Patent No. 6416959
 ; GENERAL INFORMATION:
 ; APPLICANT: Giuliano, Kenneth A.
 ; APPLICANT: Kapur, Ravi
 ; TITLE OF INVENTION: A System for Cell Based Screening
 ; FILE REFERENCE: 97-022-11
 ; CURRENT APPLICATION NUMBER: US/09/513,783A
 ; CURRENT FILING DATE: 2000-02-25
 ; NUMBER OF SEQ ID NOS: 180
 ; SOFTWARE: Patentin Ver. 2.0
 ; SEQ ID NO 178
 ; LENGTH: 805
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: GFP-NFKB
 ; Patent No. 6416959
 US-09-513-783A-178

Query Match 86.0%; Score 1289.5; DB 4; Length 805;
 Best Local Similarity 92.9%; Pred. No. 1.2e-127;
 Matches 248; Conservative 2; Mismatches 16; Indels 1; Gaps 1;

```

Cy 1 MWSKGEELFTGVVPIVLELDGVDNGHKFSVSGEGBDATYGLTLTKFICTTGKLPVPMPT 60
    |||
Db 1 MWSKGEELFTGVVPIVLELDGVDNGHKFSVSGEGBDATYGLTLTKFICTTGKLPVPMPT 60

Cy 61 LVTTLYGVOCFSRYPDHMKQHDFFKSAMPEGVVOERTIFFKDDGNVKTAEVKEGDTL 120
    |||
Db 61 LVTTLYGVOCFSRYPDHMKQHDFFKSAMPEGVVOERTIFFKDDGNVKTAEVKEGDTL 120

Cy 121 VNRLEKIDPKEDGNILGHKLEYNNSHNVYIMADKQKGIKVNFKIRHNIEDGSVOLA 180
    |||
Db 121 VNRLEKIDPKEDGNILGHKLEYNNSHNVYIMADKQKGIKVNFKIRHNIEDGSVOLA 180

Cy 181 DHYQONTPIGDGPVLPDNNHYLSTOSALSKDPNEKRDMVLEFVTAAGITLGMDELYKK 240
    |||
Db 181 DHYQONTPIGDGPVLPDNNHYLSTOSALSKDPNEKRDMVLEFVTAAGITLGMDELYKS 240

Cy 241 LSHGFPPEVEBODDGLTPMSCAQSQSGMDRHPAACAASARINV 286
    |||
Db 241 GLRSRDPPEFVDELFPLIFPAEPAQAAG 267
  
```

RESULT 4
 US-09-417-197-131
 ; Sequence 131, Application US/09417197
 ; Patent No. 6518021
 ; GENERAL INFORMATION:
 ; APPLICANT: Ole THASTRUP, et al.
 ; TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An Ir
 ; TITLE OF INVENTION: On A Cellular Response
 ; FILE REFERENCE: 3759-0110P
 ; CURRENT APPLICATION NUMBER: US/09/417,197
 ; CURRENT FILING DATE: 1999-10-07
 ; NUMBER OF SEQ ID NOS: 143
 ; SOFTWARE: Patentin version 3.0
 ; SEQ ID NO 131
 ; LENGTH: 1171
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: EGFP-NFAT fusion
 ; Patent No. 6518021
 US-09-417-197-131

Query Match 85.4%; Score 1280.5; DB 4; Length 1171;
 Best Local Similarity 93.9%; Pred. No. 1.9e-126;
 Matches 245; Conservative 1; Mismatches 6; Indels 9; Gaps 2;

```

Cy 1 MWSKGEELFTGVVPIVLELDGVDNGHKFSVSGEGBDATYGLTLTKFICTTGKLPVPMPT 60
    |||
Db 1 MWSKGEELFTGVVPIVLELDGVDNGHKFSVSGEGBDATYGLTLTKFICTTGKLPVPMPT 60

Cy 61 LVTTLYGVOCFSRYPDHMKQHDFFKSAMPEGVVOERTIFFKDDGNVKTAEVKEGDTL 120
    |||
Db 61 LVTTLYGVOCFSRYPDHMKQHDFFKSAMPEGVVOERTIFFKDDGNVKTAEVKEGDTL 120

Cy 121 VNRLEKIDPKEDGNILGHKLEYNNSHNVYIMADKQKGIKVNFKIRHNIEDGSVOLA 180
    |||
Db 121 VNRLEKIDPKEDGNILGHKLEYNNSHNVYIMADKQKGIKVNFKIRHNIEDGSVOLA 180

Cy 181 DHYQONTPIGDGPVLPDNNHYLSTOSALSKDPNEKRDMVLEFVTAAGITLGMDELYKK 240
    |||
Db 181 DHYQONTPIGDGPVLPDNNHYLSTOSALSKDPNEKRDMVLEFVTAAGITLGMDELYKS 240

Cy 241 LSHGFPPEVEBODDGLTPMSCAQSQSGMDRHPAACAASARINV 255
    |||
Db 241 GLRSRDPPEFVDELFPLIFPAEPAQAAG 258
  
```

RESULT 5
 US-09-417-197-121
 ; Sequence 121, Application US/09417197
 ; Patent No. 6518021
 ; GENERAL INFORMATION:
 ; APPLICANT: Ole THASTRUP, et al.

TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An I
FILE REFERENCE: 3759-0110P
CURRENT APPLICATION NUMBER: US/09/417,197
CURRENT FILING DATE: 1999-10-07
NUMBER OF SEQ ID NOS: 143
SOFTWARE: PatentIn version 3.0
SEQ ID NO 121
LENGTH: 997
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: EGFP-lkappab-kinase fusion
US-09-417-197-121

Query Match 85.3%; Score 1279; DB 4; Length 997;
Best Local Similarity 90.4%; Pred. No. 2,2e-126;
Matches 245; Conservative 2; Mismatches 10; Indels 14; Gaps 1;

QY 1 MWSKGEELFTGVVILVELDGVNGHKFSVSGEGGATYGLTKLFTCTTGKLPVMPPT 60
Db 1 MWSKGEELFTGVVILVELDGVNGHKFSVSGEGGATYGLTKLFTCTTGKLPVMPPT 60
QY 61 LVTLTYGVOCFSRYPDHMKQHDFFKSAMPEGYOERTIFFKDDGNKTRAVYFEGDTL 120
Db 61 LVTLTYGVOCFSRYPDHMKQHDFFKSAMPEGYOERTIFFKDDGNKTRAVYFEGDTL 120
QY 121 VNRLEKIDPKEDGNILGHKLRYNNSHNYIMADKQKGIKYNFKIRHNIEDGSVOLA 180
Db 121 VNRLEKIDPKEDGNILGHKLRYNNSHNYIMADKQKGIKYNFKIRHNIEDGSVOLA 180
QY 181 DHYOQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDMVLLFEVTAAGITLGMDELYK 240
Db 181 DHYOQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDMVLLFEVTAAGITLGMDELYK- 239
QY 241 LSHGFPPEVEBDDGTLPMSCAQESGMDRHP 271
Db 240 -----SLRSRAQAQNSMTMERPP 257

RESULT 6

US-09-417-197-47
Sequence 47, Application US/09417197
Patent No. 6518021
GENERAL INFORMATION:
APPLICANT: Ole THASTRUP, et al.
TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An I
FILE REFERENCE: 3759-0110P
CURRENT APPLICATION NUMBER: US/09/417,197
CURRENT FILING DATE: 1999-10-07
NUMBER OF SEQ ID NOS: 143
SOFTWARE: PatentIn version 3.0
SEQ ID NO 47
LENGTH: 607
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: EGFP-p38 fusion
US-09-417-197-47

Query Match 85.2%; Score 1278; DB 4; Length 607;
Best Local Similarity 84.7%; Pred. No. 1,3e-126;
Matches 249; Conservative 6; Mismatches 15; Indels 24; Gaps 3;

QY 1 MWSKGEELFTGVVILVELDGVNGHKFSVSGEGGATYGLTKLFTCTTGKLPVMPPT 60
Db 1 MWSKGEELFTGVVILVELDGVNGHKFSVSGEGGATYGLTKLFTCTTGKLPVMPPT 60
QY 61 LVTLTYGVOCFSRYPDHMKQHDFFKSAMPEGYOERTIFFKDDGNKTRAVYFEGDTL 120
Db 61 LVTLTYGVOCFSRYPDHMKQHDFFKSAMPEGYOERTIFFKDDGNKTRAVYFEGDTL 120

QY 121 VNRLEKIDPKEDGNILGHKLRYNNSHNYIMADKQKGIKYNFKIRHNIEDGSVOLA 180
Db 121 VNRLEKIDPKEDGNILGHKLRYNNSHNYIMADKQKGIKYNFKIRHNIEDGSVOLA 180
QY 181 DHYOQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDMVLLFEVTAAGITLGMDELYK- 239
Db 181 DHYOQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDMVLLFEVTAAGITLGMDELYKS 240
QY 240 -----KLSHGFPPEVEBDDGTLPMSCAQESGMDRHPAACS 276
Db 241 GLRSRAQNSMTMERPP-----GSGAYGSGVCA 288

RESULT 7

US-09-417-197-115
Sequence 115, Application US/09417197
Patent No. 6518021
GENERAL INFORMATION:
APPLICANT: Ole THASTRUP, et al.
TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An I
FILE REFERENCE: 3759-0110P
CURRENT APPLICATION NUMBER: US/09/417,197
CURRENT FILING DATE: 1999-10-07
NUMBER OF SEQ ID NOS: 143
SOFTWARE: PatentIn version 3.0
SEQ ID NO 115
LENGTH: 544
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: EGFP-CDK2 fusion
US-09-417-197-115

Query Match 85.1%; Score 1276; DB 4; Length 544;
Best Local Similarity 93.5%; Pred. No. 1,8e-126;
Matches 244; Conservative 4; Mismatches 7; Indels 6; Gaps 2;

QY 1 MWSKGEELFTGVVILVELDGVNGHKFSVSGEGGATYGLTKLFTCTTGKLPVMPPT 60
Db 1 MWSKGEELFTGVVILVELDGVNGHKFSVSGEGGATYGLTKLFTCTTGKLPVMPPT 60
QY 61 LVTLTYGVOCFSRYPDHMKQHDFFKSAMPEGYOERTIFFKDDGNKTRAVYFEGDTL 120
Db 61 LVTLTYGVOCFSRYPDHMKQHDFFKSAMPEGYOERTIFFKDDGNKTRAVYFEGDTL 120
QY 121 VNRLEKIDPKEDGNILGHKLRYNNSHNYIMADKQKGIKYNFKIRHNIEDGSVOLA 180
Db 121 VNRLEKIDPKEDGNILGHKLRYNNSHNYIMADKQKGIKYNFKIRHNIEDGSVOLA 180
QY 181 DHYOQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDMVLLFEVTAAGITLGMDELYK- 239
Db 181 DHYOQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDMVLLFEVTAAGITLGMDELYKS 240
QY 240 -----KLSHGFPPEVEBDDGT 256
Db 241 GLRSRAQNSMTMERPP-----GSGAYGSGVCA 288

RESULT 8

US-09-417-197-39
Sequence 39, Application US/09417197
Patent No. 6518021
GENERAL INFORMATION:
APPLICANT: Ole THASTRUP, et al.
TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An I
FILE REFERENCE: 3759-0110P
CURRENT APPLICATION NUMBER: US/09/417,197
CURRENT FILING DATE: 1999-10-07
NUMBER OF SEQ ID NOS: 143
SOFTWARE: PatentIn version 3.0
SEQ ID NO 39

LENGTH: 631
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: EGFP-Etk1 fusion construct
US-09-417-197-39

Query Match 85.1%; Score 1276; DB 4; Length 631;
Best Local Similarity 91.7%; Pred. No. 2.2e-126;
Matches 244; Conservative 4; Mismatches 14; Indels 4; Gaps 2;

QY 1 MWSKGBELFTGVVPIVLELDGVNGHKFSVSGEGDATTGKLTIKFTCTTGKLPVWPT 60
DB 1 MWSKGBELFTGVVPIVLELDGVNGHKFSVSGEGDATTGKLTIKFTCTTGKLPVWPT 60
QY 61 LVTLITLVGVCFSRYPDHMKQHDFFKSAMPEGYOERTIFFKDDGNKYKTRAEVKEGDTL 120
DB 61 LVTLITLVGVCFSRYPDHMKQHDFFKSAMPEGYOERTIFFKDDGNKYKTRAEVKEGDTL 120
QY 121 VNRLEKIGIDPEKEDGNILGHKLEYNNSHNYIMADKQNGIKVNFIRHNIEDGSVOLA 180
DB 121 VNRLEKIGIDPEKEDGNILGHKLEYNNSHNYIMADKQNGIKVNFIRHNIEDGSVOLA 180
QY 181 DHYOQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDMVLLFEVTAAGITLGMDELYKK 240
DB 181 DHYOQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDMVLLFEVTAAGITLGMDELYKS 240
QY 241 LSHGPPPEVERODDGLTLMSCAQBSG 266
DB 241 --GHSRRA-QASNSTMAAAAGGG 262

RESULT 9

US-09-417-197-117
Sequence 117; Application US/09417197
Patent No. 6518021
GENERAL INFORMATION:
APPLICANT: Ole Thastrup, et al.
TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An I
FILE REFERENCE: 3759-0110P
CURRENT APPLICATION NUMBER: US/09/417,197
CURRENT FILING DATE: 1999-10-07
NUMBER OF SEQ ID NOS: 143
SOFTWARE: PatentIn version 3.0
SEQ ID NO 117
LENGTH: 843
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: EGFP-PTP fusion
US-09-417-197-117

Query Match 85.1%; Score 1276; DB 4; Length 843;
Best Local Similarity 92.8%; Pred. No. 3.5e-126;
Matches 246; Conservative 2; Mismatches 9; Indels 8; Gaps 2;

QY 1 MWSKGBELFTGVVPIVLELDGVNGHKFSVSGEGDATTGKLTIKFTCTTGKLPVWPT 60
DB 1 MWSKGBELFTGVVPIVLELDGVNGHKFSVSGEGDATTGKLTIKFTCTTGKLPVWPT 60
QY 61 LVTLITLVGVCFSRYPDHMKQHDFFKSAMPEGYOERTIFFKDDGNKYKTRAEVKEGDTL 120
DB 61 LVTLITLVGVCFSRYPDHMKQHDFFKSAMPEGYOERTIFFKDDGNKYKTRAEVKEGDTL 120
QY 121 VNRLEKIGIDPEKEDGNILGHKLEYNNSHNYIMADKQNGIKVNFIRHNIEDGSVOLA 180
DB 121 VNRLEKIGIDPEKEDGNILGHKLEYNNSHNYIMADKQNGIKVNFIRHNIEDGSVOLA 180
QY 181 DHYOQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDMVLLFEVTAAGITLGMDELYKK 240
DB 181 DHYOQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDMVLLFEVTAAGITLGMDELYKS 240

QY 241 -----LSHG-PPPEVERODDGLT 257
DB 241 GLASREMLSRGWFHRLDLSGLDATT 265

RESULT 10

US-09-417-197-135
Sequence 135; Application US/09417197
Patent No. 6518021
GENERAL INFORMATION:
APPLICANT: Ole Thastrup, et al.
TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An I
FILE REFERENCE: 3759-0110P
CURRENT APPLICATION NUMBER: US/09/417,197
CURRENT FILING DATE: 1999-10-07
NUMBER OF SEQ ID NOS: 143
SOFTWARE: PatentIn version 3.0
SEQ ID NO 135
LENGTH: 933
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: EGFP-PKG fusion
US-09-417-197-135

Query Match 85.0%; Score 1275.5; DB 4; Length 933;
Best Local Similarity 91.7%; Pred. No. 4.6e-126;
Matches 242; Conservative 3; Mismatches 6; Indels 13; Gaps 1;

QY 1 MWSKGBELFTGVVPIVLELDGVNGHKFSVSGEGDATTGKLTIKFTCTTGKLPVWPT 60
DB 1 MWSKGBELFTGVVPIVLELDGVNGHKFSVSGEGDATTGKLTIKFTCTTGKLPVWPT 60
QY 61 LVTLITLVGVCFSRYPDHMKQHDFFKSAMPEGYOERTIFFKDDGNKYKTRAEVKEGDTL 120
DB 61 LVTLITLVGVCFSRYPDHMKQHDFFKSAMPEGYOERTIFFKDDGNKYKTRAEVKEGDTL 120
QY 121 VNRLEKIGIDPEKEDGNILGHKLEYNNSHNYIMADKQNGIKVNFIRHNIEDGSVOLA 180
DB 121 VNRLEKIGIDPEKEDGNILGHKLEYNNSHNYIMADKQNGIKVNFIRHNIEDGSVOLA 180
QY 181 DHYOQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDMVLLFEVTAAGITLGMDELYKK 240
DB 181 DHYOQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDMVLLFEVTAAGITLGMDELYKS 240
QY 241 -----LSHGPPPEVER 251
DB 241 GLSRGSMGTIRDLQYALQEXIEE 264

RESULT 11

US-09-513-783A-172
Sequence 172; Application US/09513783A
Patent No. 6416959
GENERAL INFORMATION:
APPLICANT: Giuliano, Kenneth A.
TITLE OF INVENTION: A System for Cell Based Screening
FILE REFERENCE: 97-022-11
CURRENT APPLICATION NUMBER: US/09/513,783A
CURRENT FILING DATE: 2000-02-25
NUMBER OF SEQ ID NOS: 180
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 172
LENGTH: 941
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: GFP-HSP70
US-09-513-783A-172

Query Match 85.0%; Score 1275.5; DB 4; Length 941;

Best Local Similarity 85.0%; Pred. No. 4.6e-126;
Matches 250; Conservative 3; Mismatches 10; Indels 31; Gaps 4;

QY 1 MWSKGEELFTGVVPIIVELDGVNGHKSFSVSGEGDGYATYKTLTKFTCTTGKLPVPMPT 60
Db 1 MWSKGEELFTGVVPIIVELDGVNGHKSFSVSGEGDGYATYKTLTKFTCTTGKLPVPMPT 60
QY 61 LVTTLVYGVCCFSRYPDHMKOHDFEKSAMPEGYVQERTIFFKDDGNYKTRAIVEKFGDTL 120
Db 61 LVTTLVYGVCCFSRYPDHMKOHDFEKSAMPEGYVQERTIFFKDDGNYKTRAIVEKFGDTL 120
QY 121 VNRLEKGIIDFKEDGNILGHKLEYNNSHNHYIMADKOKNGIKVNFKIRHNIIDSGVOLA 180
Db 121 VNRLEKGIIDFKEDGNILGHKLEYNNSHNHYIMADKOKNGIKVNFKIRHNIIDSGVOLA 180
QY 181 DHYQONTPIGDGPVLLPDNHYLSTOSALSKDPNEKRDHWLLEFVTAAGITLGMDELYKS 240
Db 181 DHYQONTPIGDGPVLLPDNHYLSTOSALSKDPNEKRDHWLLEFVTAAGITLGMDELYKS 240
QY 241 -----LSHGFPPEVEEODDGTLPMSG-----AOSG-----DRHPAACS 276
Db 241 GMSVGVGIDIGF-----QSCYAVARAGIETIANEYSDRCPTACIS 281

RESULT 12

US-09-417-197-139
; Sequence 139, Application US/09417197
; Patent No. 6518021
; GENERAL INFORMATION:
; APPLICANT: Ole THASTRUP, et al.
; TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An I
; FILE REFERENCE: 3759-0110P
; CURRENT APPLICATION NUMBER: US/09/417,197
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: Patentin Version 3.0
; SEQ ID NO 139
; LENGTH: 727
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: EGFP-PK8 fusion
US-09-417-197-139

Query Match 85.0%; Score 1275; DB 4; Length 727;

Best Local Similarity 94.2%; Pred. No. 3.5e-126;
Matches 244; Conservative 2; Mismatches 11; Indels 2; Gaps 1;

QY 1 MWSKGEELFTGVVPIIVELDGVNGHKSFSVSGEGDGYATYKTLTKFTCTTGKLPVPMPT 60
Db 1 MWSKGEELFTGVVPIIVELDGVNGHKSFSVSGEGDGYATYKTLTKFTCTTGKLPVPMPT 60
QY 61 LVTTLVYGVCCFSRYPDHMKOHDFEKSAMPEGYVQERTIFFKDDGNYKTRAIVEKFGDTL 120
Db 61 LVTTLVYGVCCFSRYPDHMKOHDFEKSAMPEGYVQERTIFFKDDGNYKTRAIVEKFGDTL 120
QY 121 VNRLEKGIIDFKEDGNILGHKLEYNNSHNHYIMADKOKNGIKVNFKIRHNIIDSGVOLA 180
Db 121 VNRLEKGIIDFKEDGNILGHKLEYNNSHNHYIMADKOKNGIKVNFKIRHNIIDSGVOLA 180
QY 181 DHYQONTPIGDGPVLLPDNHYLSTOSALSKDPNEKRDHWLLEFVTAAGITLGMDELYKS 240
Db 181 DHYQONTPIGDGPVLLPDNHYLSTOSALSKDPNEKRDHWLLEFVTAAGITLGMDELYKS 240
QY 241 -----LSHGFPPEVEEODDGTLPMSG-----AOSG-----DRHPAACS 276
Db 241 GMSVGVGIDIGF-----QSCYAVARAGIETIANEYSDRCPTACIS 281

RESULT 13
US-09-417-197-143
; Sequence 143, Application US/09417197

; Patent No. 6518021
; GENERAL INFORMATION:
; APPLICANT: Ole THASTRUP, et al.
; TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An I

; FILE REFERENCE: 3759-0110P
; CURRENT APPLICATION NUMBER: US/09/417,197
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: Patentin Version 3.0
; SEQ ID NO 143
; LENGTH: 797
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: EGFP-NFKappaB fusion
US-09-417-197-143

Query Match 85.0%; Score 1275; DB 4; Length 797;
Best Local Similarity 90.4%; Pred. No. 4.1e-126;
Matches 246; Conservative 1; Mismatches 11; Indels 14; Gaps 2;

QY 1 MWSKGEELFTGVVPIIVELDGVNGHKSFSVSGEGDGYATYKTLTKFTCTTGKLPVPMPT 60
Db 1 MWSKGEELFTGVVPIIVELDGVNGHKSFSVSGEGDGYATYKTLTKFTCTTGKLPVPMPT 60
QY 61 LVTTLVYGVCCFSRYPDHMKOHDFEKSAMPEGYVQERTIFFKDDGNYKTRAIVEKFGDTL 120
Db 61 LVTTLVYGVCCFSRYPDHMKOHDFEKSAMPEGYVQERTIFFKDDGNYKTRAIVEKFGDTL 120
QY 121 VNRLEKGIIDFKEDGNILGHKLEYNNSHNHYIMADKOKNGIKVNFKIRHNIIDSGVOLA 180
Db 121 VNRLEKGIIDFKEDGNILGHKLEYNNSHNHYIMADKOKNGIKVNFKIRHNIIDSGVOLA 180
QY 181 DHYQONTPIGDGPVLLPDNHYLSTOSALSKDPNEKRDHWLLEFVTAAGITLGMDELYKS 240
Db 181 DHYQONTPIGDGPVLLPDNHYLSTOSALSKDPNEKRDHWLLEFVTAAGITLGMDELYKS 240
QY 241 -----LSHGFPPEVEEODDGTLPMSG-----AOSG-----DRHPAACS 276
Db 241 GMSVGVGIDIGF-----QSCYAVARAGIETIANEYSDRCPTACIS 281

RESULT 14

US-09-172-063-3
; Sequence 3, Application US/09172063
; Patent No. 6150176
; GENERAL INFORMATION:
; APPLICANT: Tsielen, Roger Y.
; APPLICANT: Miyawaki, Atsushi
; APPLICANT: Ilopis, Juan
; APPLICANT: Machter, Rebekka M.
; APPLICANT: Remington, S. James
; TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
; FILE REFERENCE: 07257/071001
; CURRENT APPLICATION NUMBER: US/09/172,063
; EARLIER FILING DATE: 1998-10-13
; EARLIER FILING DATE: 1998-06-09
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 239
; TYPE: PRT
; ORGANISM: Aequorea victoria
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (0)...(0)
; OTHER INFORMATION: EGFP
US-09-172-063-3

Query Match 84.9%; Score 1274; DB 3; Length 239;

Best Local Similarity 100.0%; Pred. No. 8.2e-127;
Matches 239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY      1 MWSKGELFTGVVPIIVELDGVNGHKFSVSGEGSDATYGLTKFKICTTGKLPVPMPT 60
Db      1 MWSKGELFTGVVPIIVELDGVNGHKFSVSGEGSDATYGLTKFKICTTGKLPVPMPT 60
QY      61 LVTLITGVQCFSRYPDHMKQHDFFKSAPEGYVQERTIFFKODGNKTRAVEKFEEDTL 120
Db      61 LVTLITGVQCFSRYPDHMKQHDFFKSAPEGYVQERTIFFKODGNKTRAVEKFEEDTL 120
QY      121 VNRTELKGIIDFKEDGNILGHKLEYNVNSHNYIMADKQNGIKYNFKIRHNIEDGSVOLA 180
Db      121 VNRTELKGIIDFKEDGNILGHKLEYNVNSHNYIMADKQNGIKYNFKIRHNIEDGSVOLA 180
QY      181 DHYQONTPIGDPVLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTAAGITLGMDELK 239
Db      181 DHYQONTPIGDPVLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTAAGITLGMDELK 239
```

RESULT 15

US-09-513-783A-46
; Sequence 46 Application US/09513783A
; Patent No. 6416859
; GENERAL INFORMATION:
; APPLICANT: Giuliani, Kenneth A.
; APPLICANT: Kapur, Ravi
; TITLE OF INVENTION: A System for Cell Based Screening
; FILE REFERENCE: 97-022-11
; CURRENT APPLICATION NUMBER: US/09/513.783A
; CURRENT FILING DATE: 2000-02-25
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 46
; LENGTH: 239
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: EGFP
US-09-513-783A-46

Query Match 84.9%; Score 1274; DB 4; Length 239;

Best Local Similarity 100.0%; Pred. No. 8.2e-127;

Matches 239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY      1 MWSKGELFTGVVPIIVELDGVNGHKFSVSGEGSDATYGLTKFKICTTGKLPVPMPT 60
Db      1 MWSKGELFTGVVPIIVELDGVNGHKFSVSGEGSDATYGLTKFKICTTGKLPVPMPT 60
QY      61 LVTLITGVQCFSRYPDHMKQHDFFKSAPEGYVQERTIFFKODGNKTRAVEKFEEDTL 120
Db      61 LVTLITGVQCFSRYPDHMKQHDFFKSAPEGYVQERTIFFKODGNKTRAVEKFEEDTL 120
QY      121 VNRTELKGIIDFKEDGNILGHKLEYNVNSHNYIMADKQNGIKYNFKIRHNIEDGSVOLA 180
Db      121 VNRTELKGIIDFKEDGNILGHKLEYNVNSHNYIMADKQNGIKYNFKIRHNIEDGSVOLA 180
QY      181 DHYQONTPIGDPVLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTAAGITLGMDELK 239
Db      181 DHYQONTPIGDPVLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTAAGITLGMDELK 239
```

Search completed: May 14, 2004, 08:27:50
Job time : 23 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM protein - protein search, using sw model

Run on: May 14, 2004, 08:26:49 ; Search time 570 Seconds
(Without alignments)
137.178 Million cell updates/sec

Title: US-09-931-232-1
Perfect score: 1500
Sequence: 1 MWSKGELFTGVPIVLELD.....AQSGMDRRHPACASARINV 281

Scoring table: BLOSUM62
Gapop 10.0 , Gapept 0.5

Searched: 1145568 segs, 278261457 residues

Total number of hits satisfying chosen parameters: 1145568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1: /cgn2_6/prodata/1/pubppa/US07_PUBCOMB.pep:*
2: /cgn2_6/prodata/1/pubppa/PCF_NEW_PUB.pep:*
3: /cgn2_6/prodata/1/pubppa/US06_NEW_PUB.pep:*
4: /cgn2_6/prodata/1/pubppa/US06_PUBCOMB.pep:*
5: /cgn2_6/prodata/1/pubppa/US07_NEW_PUB.pep:*
6: /cgn2_6/prodata/1/pubppa/PCFUS_PUBCOMB.pep:*
7: /cgn2_6/prodata/1/pubppa/US08_NEW_PUB.pep:*
8: /cgn2_6/prodata/1/pubppa/US08_PUBCOMB.pep:*
9: /cgn2_6/prodata/1/pubppa/US09_PUBCOMB.pep:*
10: /cgn2_6/prodata/1/pubppa/US09_PUBCOMB.pep:*
11: /cgn2_6/prodata/1/pubppa/US09C_PUBCOMB.pep:*
12: /cgn2_6/prodata/1/pubppa/US09_NEW_PUB.pep:*
13: /cgn2_6/prodata/1/pubppa/US10_PUBCOMB.pep:*
14: /cgn2_6/prodata/1/pubppa/US10C_PUBCOMB.pep:*
15: /cgn2_6/prodata/1/pubppa/US10C_PUBCOMB.pep:*
16: /cgn2_6/prodata/1/pubppa/US10_NEW_PUB.pep:*
17: /cgn2_6/prodata/1/pubppa/US60_NEW_PUB.pep:*
18: /cgn2_6/prodata/1/pubppa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1500	100.0	281	US-09-931-232-1	Sequence 1, Appl1
2	1290.5	86.0	288	US-10-314-861-37	Sequence 37, Appl1
3	1290.5	86.0	293	US-10-314-861-35	Sequence 35, Appl1
4	1290.5	86.0	295	US-10-314-861-39	Sequence 39, Appl1
5	1290.5	86.0	299	US-10-314-861-33	Sequence 33, Appl1
6	1290.5	86.0	305	US-10-314-861-31	Sequence 31, Appl1
7	1290.5	86.0	311	US-10-314-861-29	Sequence 29, Appl1
8	1290.5	86.0	324	US-10-314-861-16	Sequence 16, Appl1
9	1289.5	86.0	805	US-10-100-957A-178	Sequence 178, Appl1
10	1289.5	85.4	1171	US-10-072-036-131	Sequence 131, Appl1
11	1279	85.3	997	US-10-072-036-121	Sequence 121, Appl1
12	1278	85.2	1089	US-10-072-036-47	Sequence 47, Appl1
13	1277	85.1	1089	US-10-259-864-2	Sequence 2, Appl1
14	1276.5	85.1	1090	US-10-259-864-6	Sequence 6, Appl1
15	1276	85.1	544	US-10-072-036-115	Sequence 115, Appl1

16	1276	85.1	631	US-10-072-036-39	Sequence 39, Appl1
17	1276	85.1	843	US-10-072-036-117	Sequence 117, Appl1
18	1275.5	85.0	931	US-10-072-036-135	Sequence 135, Appl1
19	1275.5	85.0	941	US-10-100-957A-172	Sequence 172, Appl1
20	1275	85.0	727	US-10-072-036-139	Sequence 139, Appl1
21	1275	85.0	797	US-10-072-036-143	Sequence 143, Appl1
22	1274	84.9	239	US-09-920-922-2	Sequence 2, Appl1
23	1274	84.9	239	US-09-999-745-4	Sequence 4, Appl1
24	1274	84.9	239	US-09-866-538-4	Sequence 4, Appl1
25	1274	84.9	239	US-09-797-496B-2	Sequence 2, Appl1
26	1274	84.9	239	US-09-794-308-4	Sequence 4, Appl1
27	1274	84.9	239	US-09-865-291-4	Sequence 4, Appl1
28	1274	84.9	239	US-10-457-982-3	Sequence 3, Appl1
29	1274	84.9	239	US-10-121-258-13	Sequence 13, Appl1
30	1274	84.9	239	US-10-221-461-7	Sequence 7, Appl1
31	1274	84.9	239	US-10-100-957A-2	Sequence 2, Appl1
32	1274	84.9	239	US-10-177-390-2	Sequence 2, Appl1
33	1274	84.9	239	US-10-338-411-3	Sequence 3, Appl1
34	1274	84.9	239	US-10-370-570-4	Sequence 4, Appl1
35	1274	84.9	239	US-10-389-640-3	Sequence 3, Appl1
36	1274	84.9	259	US-10-314-861-11	Sequence 11, Appl1
37	1274	84.9	294	US-10-100-957A-2	Sequence 2, Appl1
38	1274	84.9	308	US-10-033-717-35	Sequence 35, Appl1
39	1274	84.9	320	US-10-338-411-11	Sequence 11, Appl1
40	1274	84.9	320	US-10-389-640-11	Sequence 11, Appl1
41	1274	84.9	323	US-10-457-982-21	Sequence 21, Appl1
42	1274	84.9	323	US-10-338-411-7	Sequence 7, Appl1
43	1274	84.9	323	US-10-338-411-13	Sequence 13, Appl1
44	1274	84.9	323	US-10-389-640-7	Sequence 7, Appl1
45	1274	84.9	323	US-10-389-640-13	Sequence 13, Appl1

ALIGNMENTS

RESULT 1
US-09-931-232-1
Sequence 1, Application US/09931232
Publication No. US20020058274A1
GENERAL INFORMATION:
APPLICANT: Li, Xiangliang
TITLE OR INVENTION: Rapidly Degrading GFP-Fusion Proteins
FILE REFERENCE: CLONING/55CON
CURRENT APPLICATION NUMBER: US/09/931,232
CURRENT FILING DATE: 2001-08-15
PRIOR APPLICATION NUMBER: 09/364,946
PRIOR FILING DATE: 1999-07-30
PRIOR APPLICATION NUMBER: 09/191,233
PRIOR FILING DATE: 1998-11-13
PRIOR APPLICATION NUMBER: 09/062,102
PRIOR FILING DATE: 1998-04-17
PRIOR APPLICATION NUMBER: 60/060,855
PRIOR FILING DATE: 1997-10-02
NUMBER OF SEQ ID NOS: 3
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 281
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURES:
OTHER INFORMATION: Sequence of the EGFP-MODC422_461 fusion
US-09-931-232-1

Query Match 100.0%; Score 1500; DB 12; Length 281;
Best Local Similarity 100.0%; Pred. No. 1.7e-144;
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1 MWSKGELFTGVPIVLELDGVDNGHGFYSGBEGDPTGKTLTKFTCTTGKLPVWPPT 60
1 MWSKGELFTGVPIVLELDGVDNGHGFYSGBEGDPTGKTLTKFTCTTGKLPVWPPT 60

QY 61 LVTLLTYGVQCFSRYPDHMKOHDFPKSAMPEGYVOERTIFFPKDGNKYKTRAEVFEEDTL 120
DB 61 LVTLLTYGVQCFSRYPDHMKOHDFPKSAMPEGYVOERTIFFPKDGNKYKTRAEVFEEDTL 120
QY 121 VNRLEIKGIDFKEDGNILGHKLEYNVNSHNYIMADKQKNGIKVNFKIRHNIEDGSVOLA 180
DB 121 VNRLEIKGIDFKEDGNILGHKLEYNVNSHNYIMADKQKNGIKVNFKIRHNIEDGSVOLA 180
QY 181 DHYQONTPIGDPVLLPDPNHVLTSTOSALSCKDPNEKRDMVLLFEVTAAGITLGMDELYKK 240
DB 181 DHYQONTPIGDPVLLPDPNHVLTSTOSALSCKDPNEKRDMVLLFEVTAAGITLGMDELYKK 240
QY 241 LSHGPPPEVEEODDGLTFMSCAQESGMDRHPAACASARINV 281
DB 241 LSHGPPPEVEEODDGLTFMSCAQESGMDRHPAACASARINV 281

RESULT 2
US-10-314-861-37
; Sequence 37, Application US/10314861
; Publication No. US20030148269A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Holzman, Thomas F.
; APPLICANT: Harlan, John Eric
; APPLICANT: Egan, David A.
; APPLICANT: Buko, Alexander M.
; APPLICANT: Solomon, Larry R.
; APPLICANT: Lador, Uri S.
; APPLICANT: Tang, Qining
; TITLE OF INVENTION: CENTRIFUGALLY-ENHANCED METHOD OF
; TITLE OF INVENTION: DETERMINING LIGAND/TARGET AFFINITY
; FILE REFERENCE: 6404 US.01
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: US/09/270,427
; PRIOR FILING DATE: 1999-03-16
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37
; LENGTH: 288
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Protein derived from plasmid pETGFP-EF3:980-1008
US-10-314-861-37

Query Match 86.0%; Score 1290.5; DB 14; Length 288;
Best Local Similarity 95.7%; Pred. No. 4,4e-123;
Matches 243; Conservative 4; Mismatches 6; Indels 1; Gaps 1;
QY 1 MWSKGEELFTGVVPIIVLELDGVDVNGHKFSVSGEGSDATYGLTKLKTCTTGKLPVPMPT 60
DB 21 MWSKGEELFTGVVPIIVLELDGVDVNGHKFSVSGEGSDATYGLTKLKTCTTGKLPVPMPT 80
QY 61 LVTLLTYGVQCFSRYPDHMKOHDFPKSAMPEGYVOERTIFFPKDGNKYKTRAEVFEEDTL 120
DB 81 LVTLLTYGVQCFSRYPDHMKOHDFPKSAMPEGYVOERTIFFPKDGNKYKTRAEVFEEDTL 140
QY 121 VNRLEIKGIDFKEDGNILGHKLEYNVNSHNYIMADKQKNGIKVNFKIRHNIEDGSVOLA 180
DB 141 VNRLEIKGIDFKEDGNILGHKLEYNVNSHNYIMADKQKNGIKVNFKIRHNIEDGSVOLA 200
QY 181 DHYQONTPIGDPVLLPDPNHVLTSTOSALSCKDPNEKRDMVLLFEVTAAGITLGMDELYKK 240
DB 201 DHYQONTPIGDPVLLPDPNHVLTSTOSALSCKDPNEKRDMVLLFEVTAAGITLGMDELYKS 260
QY 241 LSHGPPPEVEEODD 254
DB 261 -GGAGPRIEKED 273

RESULT 3
US-10-314-861-35
; Sequence 35, Application US/10314861
; Publication No. US20030148269A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Holzman, Thomas F.
; APPLICANT: Harlan, John Eric
; APPLICANT: Egan, David A.
; APPLICANT: Buko, Alexander M.
; APPLICANT: Solomon, Larry R.
; APPLICANT: Lador, Uri S.
; APPLICANT: Tang, Qining
; TITLE OF INVENTION: CENTRIFUGALLY-ENHANCED METHOD OF
; TITLE OF INVENTION: DETERMINING LIGAND/TARGET AFFINITY
; FILE REFERENCE: 6404 US.01
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: US/09/270,427
; PRIOR FILING DATE: 1999-03-16
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 35
; LENGTH: 293
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Protein derived from plasmid pETGFP-EF3:980-1013
US-10-314-861-35

Query Match 86.0%; Score 1290.5; DB 14; Length 293;
Best Local Similarity 95.7%; Pred. No. 4.5e-123;
Matches 243; Conservative 4; Mismatches 6; Indels 1; Gaps 1;
QY 1 MWSKGEELFTGVVPIIVLELDGVDVNGHKFSVSGEGSDATYGLTKLKTCTTGKLPVPMPT 60
DB 21 MWSKGEELFTGVVPIIVLELDGVDVNGHKFSVSGEGSDATYGLTKLKTCTTGKLPVPMPT 80
QY 61 LVTLLTYGVQCFSRYPDHMKOHDFPKSAMPEGYVOERTIFFPKDGNKYKTRAEVFEEDTL 120
DB 81 LVTLLTYGVQCFSRYPDHMKOHDFPKSAMPEGYVOERTIFFPKDGNKYKTRAEVFEEDTL 140
QY 121 VNRLEIKGIDFKEDGNILGHKLEYNVNSHNYIMADKQKNGIKVNFKIRHNIEDGSVOLA 180
DB 141 VNRLEIKGIDFKEDGNILGHKLEYNVNSHNYIMADKQKNGIKVNFKIRHNIEDGSVOLA 200
QY 181 DHYQONTPIGDPVLLPDPNHVLTSTOSALSCKDPNEKRDMVLLFEVTAAGITLGMDELYKK 240
DB 201 DHYQONTPIGDPVLLPDPNHVLTSTOSALSCKDPNEKRDMVLLFEVTAAGITLGMDELYKS 260
QY 241 LSHGPPPEVEEODD 254
DB 261 -GGAGPRIEKED 273
RESULT 4
US-10-314-861-39
; Sequence 39, Application US/10314861
; Publication No. US20030148269A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Holzman, Thomas F.
; APPLICANT: Harlan, John Eric
; APPLICANT: Egan, David A.
; APPLICANT: Buko, Alexander M.
; APPLICANT: Solomon, Larry R.
; APPLICANT: Lador, Uri S.
; APPLICANT: Tang, Qining
; TITLE OF INVENTION: CENTRIFUGALLY-ENHANCED METHOD OF
; TITLE OF INVENTION: DETERMINING LIGAND/TARGET AFFINITY
; FILE REFERENCE: 6404 US.01
; CURRENT FILING DATE: 2002-12-09

PRIOR APPLICATION NUMBER: US/09/270,427
PRIOR FILING DATE: 1999-03-16
NUMBER OF SEQ ID NOS: 39
SOFTWARE: FASTSEQ For Windows Version 4.0
SEQ ID NO 39
LENGTH: 295
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Protein derived from plasmid pETGFP-EP3:980-1008
US-10-314-861-39

Query Match 86.0%; Score 1290.5; DB 14; Length 295;
Best Local Similarity 95.7%; Pred. No. 4,6e-123;
Matches 243; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 1 MWSKGEELFTGVVPIIVLELDGVDNGHKFSVSGEGDGYTKLTKFTCTTGKLPVWPPT 60
DB 21 MWSKGEELFTGVVPIIVLELDGVDNGHKFSVSGEGDGYTKLTKFTCTTGKLPVWPPT 80
QY 61 LVTTLTGVQCFSRYPDHMKQDFFKSAMPEGVQERTIFFKDDGNKTKRAEVFEEDTL 120
DB 81 LVTTLTGVQCFSRYPDHMKQDFFKSAMPEGVQERTIFFKDDGNKTKRAEVFEEDTL 140
QY 121 VNRTELKIDFKEDGNILGHKLEYNNSHNHYIMADKQNGIKVNFIRHNIEDGSVOLA 180
DB 141 VNRTELKIDFKEDGNILGHKLEYNNSHNHYIMADKQNGIKVNFIRHNIEDGSVOLA 200
QY 181 DHYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHVLLEFVTAAGITLGMDELYKK 240
DB 201 DHYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHVLLEFVTAAGITLGMDELYKS 260
QY 241 LSHGFPPEVEEODD 254
DB 261 -GGAGAPRIEKKED 273

RESULT 5

US-10-314-861-33
Sequence 33, Application US/10314861
Publication No. US20030148269A1
GENERAL INFORMATION:
APPLICANT: Abbott Laboratories
APPLICANT: Holzman, Thomas F.
APPLICANT: Harlan, John Eric
APPLICANT: Egan, David A.
APPLICANT: Buko, Alexander M.
APPLICANT: Solomon, Larry R.
APPLICANT: Lador, Uri S.
APPLICANT: Tang, Qiang
TITLE OF INVENTION: CENTRIFUGALITY-ENHANCED METHOD OF
TITLE OF INVENTION: DETERMINING LIGAND/TARGET AFFINITY
FILE REFERENCE: 6404.US.O1
CURRENT APPLICATION NUMBER: US/10/314,861
CURRENT FILING DATE: 2002-12-09
PRIOR APPLICATION NUMBER: US/09/270,427
PRIOR FILING DATE: 1999-03-16
NUMBER OF SEQ ID NOS: 39
SOFTWARE: FASTSEQ For Windows Version 4.0
SEQ ID NO 33
LENGTH: 299
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Protein derived from plasmid pETGFP-EP3:980-1019
US-10-314-861-33

Query Match 86.0%; Score 1290.5; DB 14; Length 299;
Best Local Similarity 95.7%; Pred. No. 4,6e-123;
Matches 243; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 1 MWSKGEELFTGVVPIIVLELDGVDNGHKFSVSGEGDGYTKLTKFTCTTGKLPVWPPT 60
DB 21 MWSKGEELFTGVVPIIVLELDGVDNGHKFSVSGEGDGYTKLTKFTCTTGKLPVWPPT 80

DB 21 MWSKGEELFTGVVPIIVLELDGVDNGHKFSVSGEGDGYTKLTKFTCTTGKLPVWPPT 80
QY 61 LVTTLTGVQCFSRYPDHMKQDFFKSAMPEGVQERTIFFKDDGNKTKRAEVFEEDTL 120
DB 81 LVTTLTGVQCFSRYPDHMKQDFFKSAMPEGVQERTIFFKDDGNKTKRAEVFEEDTL 140
QY 121 VNRTELKIDFKEDGNILGHKLEYNNSHNHYIMADKQNGIKVNFIRHNIEDGSVOLA 180
DB 141 VNRTELKIDFKEDGNILGHKLEYNNSHNHYIMADKQNGIKVNFIRHNIEDGSVOLA 200
QY 181 DHYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHVLLEFVTAAGITLGMDELYKK 240
DB 201 DHYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHVLLEFVTAAGITLGMDELYKS 260
QY 241 LSHGFPPEVEEODD 254
DB 261 -GGAGAPRIEKKED 273

RESULT 6

US-10-314-861-31
Sequence 31, Application US/10314861
Publication No. US20030148269A1
GENERAL INFORMATION:
APPLICANT: Abbott Laboratories
APPLICANT: Holzman, Thomas F.
APPLICANT: Harlan, John Eric
APPLICANT: Egan, David A.
APPLICANT: Buko, Alexander M.
APPLICANT: Solomon, Larry R.
APPLICANT: Lador, Uri S.
APPLICANT: Tang, Qiang
TITLE OF INVENTION: CENTRIFUGALITY-ENHANCED METHOD OF
TITLE OF INVENTION: DETERMINING LIGAND/TARGET AFFINITY
FILE REFERENCE: 6404.US.O1
CURRENT APPLICATION NUMBER: US/10/314,861
CURRENT FILING DATE: 2002-12-09
PRIOR APPLICATION NUMBER: US/09/270,427
PRIOR FILING DATE: 1999-03-16
NUMBER OF SEQ ID NOS: 39
SOFTWARE: FASTSEQ For Windows Version 4.0
SEQ ID NO 31
LENGTH: 305
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Protein derived from plasmid pETGFP-EP3:980-1025
US-10-314-861-31

Query Match 86.0%; Score 1290.5; DB 14; Length 305;
Best Local Similarity 95.7%; Pred. No. 4,8e-123;
Matches 243; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 1 MWSKGEELFTGVVPIIVLELDGVDNGHKFSVSGEGDGYTKLTKFTCTTGKLPVWPPT 60
DB 21 MWSKGEELFTGVVPIIVLELDGVDNGHKFSVSGEGDGYTKLTKFTCTTGKLPVWPPT 80
QY 61 LVTTLTGVQCFSRYPDHMKQDFFKSAMPEGVQERTIFFKDDGNKTKRAEVFEEDTL 120
DB 81 LVTTLTGVQCFSRYPDHMKQDFFKSAMPEGVQERTIFFKDDGNKTKRAEVFEEDTL 140
QY 121 VNRTELKIDFKEDGNILGHKLEYNNSHNHYIMADKQNGIKVNFIRHNIEDGSVOLA 180
DB 141 VNRTELKIDFKEDGNILGHKLEYNNSHNHYIMADKQNGIKVNFIRHNIEDGSVOLA 200
QY 181 DHYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHVLLEFVTAAGITLGMDELYKK 240
DB 201 DHYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHVLLEFVTAAGITLGMDELYKS 260
QY 241 LSHGFPPEVEEODD 254
DB 261 -GGAGAPRIEKKED 273

```
RESULT 7
US-10-314-861-29
; Sequence 29, Application US/10314861
; Publication No. US20030148269A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Holzman, Thomas F.
; APPLICANT: Harlan, John Eric
; APPLICANT: Egan, David A.
; APPLICANT: Buko, Alexander M.
; APPLICANT: Solomon, Larry R.
; APPLICANT: Lador, Uri S.
; APPLICANT: Tang, Qiang
; TITLE OF INVENTION: CENTRIFUGALLY-ENHANCED METHOD OF
; TITLE OF INVENTION: DETERMINING LIGAND/TARGET AFFINITY
; FILE REFERENCE: 6404.US.O1
; CURRENT APPLICATION NUMBER: US/10/314,861
; PRIOR FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: US/09/270,427
; PRIOR FILING DATE: 1999-03-16
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Protein derived from plasmid pETGFP-BF3:980-1031
US-10-314-861-29

Query Match      86.0%; Score 1290.5; DB 14; Length 311;
Best Local Similarity 95.7%; Pred. No. 4,9e-123;
Matches 243; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 1 MWSKGEELFTGVVPLIVELDGVNGHKFSVSGEGBDATYGLTKLFTCTTGKLPVWPPT 60
DB 21 MWSKGEELFTGVVPLIVELDGVNGHKFSVSGEGBDATYGLTKLFTCTTGKLPVWPPT 80
QY 61 LVTTLLTYGVQCFSRYPDHMKQDFFKSAMPEGVQERTIFFKDDGNVKTAAVKEFGDTL 120
DB 81 LVTTLLTYGVQCFSRYPDHMKQDFFKSAMPEGVQERTIFFKDDGNVKTAAVKEFGDTL 140
QY 121 VNRIELKGDIFKEDGNILGHKLEYNVNSHNYIMADKONGIKVNFKIRHNIEDGSVOLA 180
DB 141 VNRIELKGDIFKEDGNILGHKLEYNVNSHNYIMADKONGIKVNFKIRHNIEDGSVOLA 200
QY 181 DHVQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHNVLEFYTAAGITLGMDELYKS 240
DB 201 DHVQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHNVLEFYTAAGITLGMDELYKS 260
QY 241 LSHGFPPEVEODD 254
DB 261 -GGAGGPRIEKED 273

RESULT 8
US-10-314-861-16
; Sequence 16, Application US/10314861
; Publication No. US20030148269A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Holzman, Thomas F.
; APPLICANT: Harlan, John Eric
; APPLICANT: Egan, David A.
; APPLICANT: Buko, Alexander M.
; APPLICANT: Solomon, Larry R.
; APPLICANT: Lador, Uri S.
; APPLICANT: Tang, Qiang
; TITLE OF INVENTION: CENTRIFUGALLY-ENHANCED METHOD OF
; TITLE OF INVENTION: DETERMINING LIGAND/TARGET AFFINITY
; FILE REFERENCE: 6404.US.O1
; CURRENT APPLICATION NUMBER: US/10/314,861
```

```
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: US/09/270,427
; PRIOR FILING DATE: 1999-03-16
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 324
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Protein derived from plasmid pETGFP-BF3:980-1044
US-10-314-861-16

Query Match      86.0%; Score 1290.5; DB 14; Length 324;
Best Local Similarity 95.7%; Pred. No. 5,2e-123;
Matches 243; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 1 MWSKGEELFTGVVPLIVELDGVNGHKFSVSGEGBDATYGLTKLFTCTTGKLPVWPPT 60
DB 21 MWSKGEELFTGVVPLIVELDGVNGHKFSVSGEGBDATYGLTKLFTCTTGKLPVWPPT 80
QY 61 LVTTLLTYGVQCFSRYPDHMKQDFFKSAMPEGVQERTIFFKDDGNVKTAAVKEFGDTL 120
DB 81 LVTTLLTYGVQCFSRYPDHMKQDFFKSAMPEGVQERTIFFKDDGNVKTAAVKEFGDTL 140
QY 121 VNRIELKGDIFKEDGNILGHKLEYNVNSHNYIMADKONGIKVNFKIRHNIEDGSVOLA 180
DB 141 VNRIELKGDIFKEDGNILGHKLEYNVNSHNYIMADKONGIKVNFKIRHNIEDGSVOLA 200
QY 181 DHVQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHNVLEFYTAAGITLGMDELYKS 240
DB 201 DHVQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHNVLEFYTAAGITLGMDELYKS 260
QY 241 LSHGFPPEVEODD 254
DB 261 -GGAGGPRIEKED 273

RESULT 9
US-10-100-957A-178
; Sequence 178, Application US/10100957A
; Publication No. US20030096322A1
; GENERAL INFORMATION:
; APPLICANT: Giuliani, Kenneth A.
; APPLICANT: Kapur, Ravi
; TITLE OF INVENTION: A System for Cell Based Screening
; FILE REFERENCE: 97-022-11A
; CURRENT APPLICATION NUMBER: US/10/100,957A
; CURRENT FILING DATE: 2002-03-19
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 178
; LENGTH: 805
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GFP-NFKB
US-10-100-957A-178

Query Match      86.0%; Score 1289.5; DB 14; Length 805;
Best Local Similarity 92.9%; Pred. No. 2,5e-122;
Matches 248; Conservative 2; Mismatches 16; Indels 1; Gaps 1;

QY 1 MWSKGEELFTGVVPLIVELDGVNGHKFSVSGEGBDATYGLTKLFTCTTGKLPVWPPT 60
DB 1 MWSKGEELFTGVVPLIVELDGVNGHKFSVSGEGBDATYGLTKLFTCTTGKLPVWPPT 60
QY 61 LVTTLLTYGVQCFSRYPDHMKQDFFKSAMPEGVQERTIFFKDDGNVKTAAVKEFGDTL 120
DB 61 LVTTLLTYGVQCFSRYPDHMKQDFFKSAMPEGVQERTIFFKDDGNVKTAAVKEFGDTL 120
QY 121 VNRIELKGDIFKEDGNILGHKLEYNVNSHNYIMADKONGIKVNFKIRHNIEDGSVOLA 180
```

Db 121 VNRLEKIDFKEDGNILGHKLEYNNSHNVYIMADKQKNGIKVNFKIRNIEDGSVOA 180
Qy 181 DHYQONTPIGDGVLLPDNNHYLSTQSALSKDPNEKDHMLLEFVTAAGITLGMDELYK 240
Db 181 DHYQONTPIGDGVLLPDNNHYLSTQSALSKDPNEKDHMLLEFVTAAGITLGMDELYK 240
Qy 241 -LSHGFPPEVEODDGTLPMSCAOESG 266
Db 241 GLRSRDPPEWDELFLIFPAAPQASG 267

RESULT 10

US-10-072-036-131
; Sequence 131, Application US/10072036
; Publication No. US20030082564A1
; GENERAL INFORMATION:
; APPLICANT: Ole THASTRUP
; APPLICANT: Sara BJORN
; APPLICANT: Soren TULLIN
; APPLICANT: Kasper ALMHOLT
; APPLICANT: Kurt SCUDDER
; TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An I
; FILE REFERENCE: 3759-0120P
; CURRENT APPLICATION NUMBER: US/10/072,036
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 09/417,197
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 131
; LENGTH: 1171
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: EGFP-NEAT fusion
US-10-072-036-131

Query Match 85.4%; Score 1280.5; DB 14; Length 1171;

Best Local Similarity 93.9%; Pred. No. 3.6e-121; Mismatches 6; Indels 9; Gaps 2;

Matches 245; Conservative 1; Mismatches 6; Indels 9; Gaps 2;
Qy 1 MVSKEELFTGVVPIIVELDDGVNGHKFSVSGEGSDATYKLTLEICTTGKLPVMPPT 60
Db 1 MVSKEELFTGVVPIIVELDDGVNGHKFSVSGEGSDATYKLTLEICTTGKLPVMPPT 60
Qy 61 LVTTLYGVQCFSRYPDHMKQHDFFKSAMPEGYQERTIFFKDDGNYKTRAIVKFEQDTL 120
Db 61 LVTTLYGVQCFSRYPDHMKQHDFFKSAMPEGYQERTIFFKDDGNYKTRAIVKFEQDTL 120
Qy 121 VNRLEKIDFKEDGNILGHKLEYNNSHNVYIMADKQKNGIKVNFKIRNIEDGSVOA 180
Db 121 VNRLEKIDFKEDGNILGHKLEYNNSHNVYIMADKQKNGIKVNFKIRNIEDGSVOA 180
Qy 181 DHYQONTPIGDGVLLPDNNHYLSTQSALSKDPNEKDHMLLEFVTAAGITLGMDELYK 240
Db 181 DHYQONTPIGDGVLLPDNNHYLSTQSALSKDPNEKDHMLLEFVTAAGITLGMDELYK 240
Qy 241 LSHGF-----PEVEEODDG 255
Db 241 ---GLRSRANMAPRQPDG 258

RESULT 11

US-10-072-036-121
; Sequence 121, Application US/10072036
; Publication No. US20030082564A1
; GENERAL INFORMATION:
; APPLICANT: Ole THASTRUP
; APPLICANT: Sara BJORN
; APPLICANT: Soren TULLIN
; APPLICANT: Kasper ALMHOLT
; APPLICANT: Kurt SCUDDER

; TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An I
; FILE REFERENCE: 3759-0120P
; CURRENT APPLICATION NUMBER: US/10/072,036
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 09/417,197
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 121
; LENGTH: 997
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: EGFP-lkappab-kinase fusion
US-10-072-036-121

Query Match 85.3%; Score 1279; DB 14; Length 997;

Best Local Similarity 90.4%; Pred. No. 4.1e-121; Mismatches 10; Indels 14; Gaps 1;

Matches 245; Conservative 2; Mismatches 10; Indels 14; Gaps 1;
Qy 1 MVSKEELFTGVVPIIVELDDGVNGHKFSVSGEGSDATYKLTLEICTTGKLPVMPPT 60
Db 1 MVSKEELFTGVVPIIVELDDGVNGHKFSVSGEGSDATYKLTLEICTTGKLPVMPPT 60
Qy 61 LVTTLYGVQCFSRYPDHMKQHDFFKSAMPEGYQERTIFFKDDGNYKTRAIVKFEQDTL 120
Db 61 LVTTLYGVQCFSRYPDHMKQHDFFKSAMPEGYQERTIFFKDDGNYKTRAIVKFEQDTL 120
Qy 121 VNRLEKIDFKEDGNILGHKLEYNNSHNVYIMADKQKNGIKVNFKIRNIEDGSVOA 180
Db 121 VNRLEKIDFKEDGNILGHKLEYNNSHNVYIMADKQKNGIKVNFKIRNIEDGSVOA 180
Qy 181 DHYQONTPIGDGVLLPDNNHYLSTQSALSKDPNEKDHMLLEFVTAAGITLGMDELYK 240
Db 181 DHYQONTPIGDGVLLPDNNHYLSTQSALSKDPNEKDHMLLEFVTAAGITLGMDELYK- 239
Qy 241 LSHGFPPEVEODDGTLPMSCAOESGMDRHP 271
Db 240 -----GLRSRANMAPRQPDG 257

RESULT 12

US-10-072-036-47
; Sequence 47, Application US/10072036
; Publication No. US20030082564A1
; GENERAL INFORMATION:
; APPLICANT: Ole THASTRUP
; APPLICANT: Sara BJORN
; APPLICANT: Soren TULLIN
; APPLICANT: Kasper ALMHOLT
; APPLICANT: Kurt SCUDDER
; TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An I
; FILE REFERENCE: 3759-0120P
; CURRENT APPLICATION NUMBER: US/10/072,036
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 09/417,197
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 47
; LENGTH: 607
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: EGFP-p38 fusion
US-10-072-036-47

Query Match 85.2%; Score 1278; DB 14; Length 607;
Best Local Similarity 84.7%; Pred. No. 2.5e-121;
Matches 249; Conservative 6; Mismatches 15; Indels 24; Gaps 3;

```
QY 1 MWSKGEELFTGVVPLIVELDGVNGHKFSVSGEGEDATYGLTLKFTCTTGKLPVPMPT 60
DB 1 MWSKGEELFTGVVPLIVELDGVNGHKFSVSGEGEDATYGLTLKFTCTTGKLPVPMPT 60
QY 61 LVTLTYGVQCFSSRYPDHMKOHDFPKSAMPEGYQERTIFFKODGNKTRAEVFEEDTL 120
DB 61 LVTLTYGVQCFSSRYPDHMKOHDFPKSAMPEGYQERTIFFKODGNKTRAEVFEEDTL 120
QY 121 VNRLEKIDPEKEDGNITLGHKLEYNNSHNYIMADKOKNGIKVNFIRHNIEDGSYOLA 180
DB 121 VNRLEKIDPEKEDGNITLGHKLEYNNSHNYIMADKOKNGIKVNFIRHNIEDGSYOLA 180
QY 181 DHYOQNTPIGDGPVLLPNDNHYLSTQSALSKDPNEKRDMVLLFEVTAAGITLGMDELKX 239
DB 181 DHYOQNTPIGDGPVLLPNDNHYLSTQSALSKDPNEKRDMVLLFEVTAAGITLGMDELKX 240
QY 240 -----LSHGPPPEVEEDDGTLPMSCAQESGMDRHPAACAS 276
DB 241 GHSRGRKMSQERPFYRQELNKTWEVERYQNLSPV-----GSGAYGSVCAA 288
```

```
RESULT 13
US-10-259-864-2
; Sequence 2, Application US/10259864
; Publication No. US2003007645A1
; GENERAL INFORMATION:
; APPLICANT: Hager, Gordon L
; TITLE OF INVENTION: Superfamily Receptor Chimeras, Translocation Assay For Superfamily
; TITLE OF INVENTION: Receptor Ligands, and Methods and Kits For Detecting and Character
; TITLE OF INVENTION: Ligands
; FILE REFERENCE: 24799
; CURRENT APPLICATION NUMBER: US/10/259,864
; PRIOR FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: US 60/325,178
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1089
; TYPE: PRT
; ORGANISM: Rat/human chimera
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (1)..()
; OTHER INFORMATION: Chimeric Protein
US-10-259-864-2
```

```
Query Match 85.1%; Score 1277; DB 14; Length 1089;
Best Local Similarity 86.4%; Pred. No. 7,4e-121;
Matches 247; Conservative 8; Mismatches 25; Indels 6; Gaps 2;
```

```
QY 1 MWSKGEELFTGVVPLIVELDGVNGHKFSVSGEGEDATYGLTLKFTCTTGKLPVPMPT 60
DB 1 MWSKGEELFTGVVPLIVELDGVNGHKFSVSGEGEDATYGLTLKFTCTTGKLPVPMPT 60
QY 61 LVTLTYGVQCFSSRYPDHMKOHDFPKSAMPEGYQERTIFFKODGNKTRAEVFEEDTL 120
DB 61 LVTLTYGVQCFSSRYPDHMKOHDFPKSAMPEGYQERTIFFKODGNKTRAEVFEEDTL 120
QY 121 VNRLEKIDPEKEDGNITLGHKLEYNNSHNYIMADKOKNGIKVNFIRHNIEDGSYOLA 180
DB 121 VNRLEKIDPEKEDGNITLGHKLEYNNSHNYIMADKOKNGIKVNFIRHNIEDGSYOLA 180
QY 181 DHYOQNTPIGDGPVLLPNDNHYLSTQSALSKDPNEKRDMVLLFEVTAAGITLGMDELKX 240
DB 181 DHYOQNTPIGDGPVLLPNDNHYLSTQSALSKDPNEKRDMVLLFEVTAAGITLGMDELKX 240
QY 241 -----LSHGPPPEVEEDDGTLPMSCAQESGMDRHPAACASARINV 281
DB 241 GALILDSKESLAPPGRDEVPGL-LGGGRGSVMDFYSLRGATVKV 285
```

RESULT 14

```
US-10-259-864-6
; Sequence 6, Application US/10259864
; Publication No. US2003007645A1
; GENERAL INFORMATION:
; APPLICANT: Hager, Gordon L
; TITLE OF INVENTION: Superfamily Receptor Chimeras, Translocation Assay For Superfamily
; TITLE OF INVENTION: Receptor Ligands, and Methods and Kits For Detecting and Character
; TITLE OF INVENTION: Ligands
; FILE REFERENCE: 24799
; CURRENT APPLICATION NUMBER: US/10/259,864
; PRIOR FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: US 60/325,178
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 1090
; TYPE: PRT
; ORGANISM: Rat/human Chimera
US-10-259-864-6
```

```
Query Match 85.1%; Score 1276.5; DB 14; Length 1090;
Best Local Similarity 86.1%; Pred. No. 8,4e-121;
Matches 247; Conservative 8; Mismatches 25; Indels 7; Gaps 2;
```

```
QY 1 MWSKGEELFTGVVPLIVELDGVNGHKFSVSGEGEDATYGLTLKFTCTTGKLPVPMPT 60
DB 1 MWSKGEELFTGVVPLIVELDGVNGHKFSVSGEGEDATYGLTLKFTCTTGKLPVPMPT 60
QY 61 LVTLTYGVQCFSSRYPDHMKOHDFPKSAMPEGYQERTIFFKODGNKTRAEVFEEDTL 120
DB 61 LVTLTYGVQCFSSRYPDHMKOHDFPKSAMPEGYQERTIFFKODGNKTRAEVFEEDTL 120
QY 121 VNRLEKIDPEKEDGNITLGHKLEYNNSHNYIMADKOKNGIKVNFIRHNIEDGSYOLA 180
DB 121 VNRLEKIDPEKEDGNITLGHKLEYNNSHNYIMADKOKNGIKVNFIRHNIEDGSYOLA 180
QY 181 DHYOQNTPIGDGPVLLPNDNHYLSTQSALSKDPNEKRDMVLLFEVTAAGITLGMDELKX 240
DB 181 DHYOQNTPIGDGPVLLPNDNHYLSTQSALSKDPNEKRDMVLLFEVTAAGITLGMDELKX 240
QY 241 -----LSHGPPPEVEEDDGTLPMSCAQESGMDRHPAACASARINV 281
DB 241 GALILDSKESLAPPGRDEVPGL-LGGGRGSVMDFYSLRGATVKV 286
```

```
RESULT 15
US-10-072-036-115
; Sequence 115, Application US/10072036
; Publication No. US20030082564A1
; GENERAL INFORMATION:
; APPLICANT: Ole THASTRUP
; APPLICANT: Sara BJORN
; APPLICANT: Soren TUDLIN
; APPLICANT: Kasper ALMHOLT
; TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An It
; TITLE OF INVENTION: On A Cellular Response
; FILE REFERENCE: 3759-0120P
; CURRENT APPLICATION NUMBER: US/10/072,036
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 09/417,197
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 115
; LENGTH: 544
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: EGFP-CDK2 fusion
US-10-072-036-115
```

Query Match 85.1%; Score 1276; DB 14; Length 544;
 Best Local Similarity 93.5%; Pred. No. 3.4e-121;
 Matches 244; Conservative 4; Mismatches 7; Indels 6; Gaps 2;

QY	1	MVSKGEELFTGVVPILVELDGDVNGHKFSVS	GGEGDATTGKLT	TKFICTTGKLPVP	PT	60
Db	1	MVSKGEELFTGVVPILVELDGDVNGHKFSVS	GGEGDATTGKLT	TKFICTTGKLPVP	PT	60
QY	61	LVTTLTGVQCFSRYPDMKQHDFFKSAMPRGY	VOERTTFPKDDGNYKTRAEVKFEGDTL			120
Db	61	LVTTLTGVQCFSRYPDMKQHDFFKSAMPRGY	VOERTTFPKDDGNYKTRAEVKFEGDTL			120
QY	121	VNRTELKGIIDFKEDGNILGHKLEYNNSHNYI	IMADKQKNGIKVNFKIRHNIEDGSVOLA			180
Db	121	VNRTELKGIIDFKEDGNILGHKLEYNNSHNYI	IMADKQKNGIKVNFKIRHNIEDGSVOLA			180
QY	181	DHYQONTPIIDGPVLLPDNHYLSTQSALS	KDPNEKRDMVLLFVTAAGITLGMDELYK			239
Db	181	DHYQONTPIIDGPVLLPDNHYLSTQSALS	KDPNEKRDMVLLFVTAAGITLGMDELYK			240
QY	240	----KLGHGPPPEYBEQDDGT				256
Db	241	GLRSRAMENF-QKVEKIGEGT				260

Search completed: May 14, 2004, 08:41:20
 Job time : 570 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 14, 2004, 08:25:19 ; Search time 183 Seconds

(Without alignments)
1498.743 Million cell updates/sec

Title: US-09-931-232-1

Perfect score: 1500

Sequence: 1 MWKGEHLFGVPIVLELD.....AQSGMDRHPACASARINV 281

Scoring table: BLOSUM62

Gapop 10.0 , Gapevt 0.5

Searched: 6019581 seqs, 976053577 residues

Total number of hits satisfying chosen parameters: 6019581

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents AA Main:*

1: /cgn2_6/prodata/2/paa/PCTUS_COMB.pep:*

2: /cgn2_6/prodata/2/paa/US06_COMB.pep:*

3: /cgn2_6/prodata/2/paa/US07_COMB.pep:*

4: /cgn2_6/prodata/2/paa/US080_COMB.pep:*

5: /cgn2_6/prodata/2/paa/US081_COMB.pep:*

6: /cgn2_6/prodata/2/paa/US082_COMB.pep:*

7: /cgn2_6/prodata/2/paa/US083_COMB.pep:*

8: /cgn2_6/prodata/2/paa/US084_COMB.pep:*

9: /cgn2_6/prodata/2/paa/US085_COMB.pep:*

10: /cgn2_6/prodata/2/paa/US086_COMB.pep:*

11: /cgn2_6/prodata/2/paa/US087_COMB.pep:*

12: /cgn2_6/prodata/2/paa/US088_COMB.pep:*

13: /cgn2_6/prodata/2/paa/US089_COMB.pep:*

14: /cgn2_6/prodata/2/paa/US090_COMB.pep:*

15: /cgn2_6/prodata/2/paa/US091_COMB.pep:*

16: /cgn2_6/prodata/2/paa/US092_COMB.pep:*

17: /cgn2_6/prodata/2/paa/US093_COMB.pep:*

18: /cgn2_6/prodata/2/paa/US094_COMB.pep:*

19: /cgn2_6/prodata/2/paa/US095_COMB.pep:*

20: /cgn2_6/prodata/2/paa/US096_COMB.pep:*

21: /cgn2_6/prodata/2/paa/US097A_COMB.pep:*

22: /cgn2_6/prodata/2/paa/US097B_COMB.pep:*

23: /cgn2_6/prodata/2/paa/US098_COMB.pep:*

24: /cgn2_6/prodata/2/paa/US099A_COMB.pep:*

25: /cgn2_6/prodata/2/paa/US099B_COMB.pep:*

26: /cgn2_6/prodata/2/paa/US100_COMB.pep:*

27: /cgn2_6/prodata/2/paa/US101_COMB.pep:*

28: /cgn2_6/prodata/2/paa/US102_COMB.pep:*

29: /cgn2_6/prodata/2/paa/US103_COMB.pep:*

30: /cgn2_6/prodata/2/paa/US104_COMB.pep:*

31: /cgn2_6/prodata/2/paa/US106_COMB.pep:*

32: /cgn2_6/prodata/2/paa/US107_COMB.pep:*

33: /cgn2_6/prodata/2/paa/US60_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match Length	ID	Description
------------	-------	--------------	----	-------------

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	
1500	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	
281	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15
Sequence 1, Appl1	Sequence 1, Appl1	Sequence 1, Appl1	Sequence 1, Appl1	Sequence 1, Appl1	Sequence 1, Appl1	Sequence 1, Appl1	Sequence 1, Appl1	Sequence 1, Appl1	Sequence 1, Appl1	Sequence 1, Appl1	Sequence 1, Appl1	Sequence 1, Appl1	Sequence 1, Appl1	Sequence 1, Appl1	Sequence 1, Appl1	Sequence 1, Appl1	Sequence 1, Appl1	Sequence 1, Appl1	Sequence 1, Appl1	Sequence 1, Appl1	Sequence 1, Appl1	Sequence 1, Appl1	Sequence 1, Appl1	Sequence 1, Appl1	Sequence 1, Appl1	Sequence 1, Appl1	Sequence 1, Appl1	Sequence 1, Appl1	Sequence 1, Appl1	Sequence 1, Appl1	Sequence 1, Appl1	Sequence 1, Appl1	Sequence 1, Appl1	Sequence 1, Appl1	Sequence 1, Appl1	Sequence 1, Appl1	Sequence 1, Appl1	Sequence 1, Appl1	Sequence 1, Appl1	Sequence 1, Appl1	Sequence 1, Appl1	Sequence 1, Appl1	Sequence 1, Appl1	Sequence 1, Appl1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1

PCT-US98-24323-1

Sequence 1, Application PC/TUS9824323A

GENERAL INFORMATION:

APPLICANT: Kain, Steve

APPLICANT: Li, Xiangyang

TITLE OF INVENTION: Rapidly Degrading GFP-Fusion Proteins and Methods

TITLE OF INVENTION: of use

FILE REFERENCE: D6100PCT

CURRENT APPLICATION NUMBER: PCT/US98/24323A

EARLIER FILING DATE: 1998-11-13

EARLIER APPLICATION NUMBER: US 09/062,102

EARLIER FILING DATE: 1998-04-17

NUMBER OF SEQ ID NOS: 3

SEQ ID NO 1

LENGTH: 281

TYPE: PRT

ORGANISM: artificial sequence

FEATURE:

OTHER INFORMATION: Sequence of the EGFP-MODC422-461 fusion protein.

PCT-US98-24323-1

Query Match 100.0%; Score 1500; DB 1; Length 281;

Best Local Similarity 100.0%; Pred. No. 1e-151;
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 1 MWSKGEELFTGVVPIVLELDGVNGHKFSVSGEGDATTYKLTIKFTCTTGKLPVPMPT 60
Db 1 MWSKGEELFTGVVPIVLELDGVNGHKFSVSGEGDATTYKLTIKFTCTTGKLPVPMPT 60
OY 61 LVTTLTLYGQCFSRPYDHNKQDFFKSNMPEGVQERTIFFKDDGNKTRAEVKEGDTL 120
Db 61 LVTTLTLYGQCFSRPYDHNKQDFFKSNMPEGVQERTIFFKDDGNKTRAEVKEGDTL 120
OY 121 VNRLEIKGIDFKEDGNILGHKLEYNVNSHNYIMADKQNGIKVNFKIRHNIEDGSVOLA 180
Db 121 VNRLEIKGIDFKEDGNILGHKLEYNVNSHNYIMADKQNGIKVNFKIRHNIEDGSVOLA 180
OY 181 DHYQONTPIGDGPVLLPDNHYLSTOSALS KDPNEKRDMVLLLEFVTAAGITLGMDELYKK 240
Db 181 DHYQONTPIGDGPVLLPDNHYLSTOSALS KDPNEKRDMVLLLEFVTAAGITLGMDELYKK 240
OY 241 LSHGPPPEVEEODDGTLPMSCAQESGMDRHPAACAASARINV 281
Db 241 LSHGPPPEVEEODDGTLPMSCAQESGMDRHPAACAASARINV 281
```

RESULT 2

US-09-191-233-1
; Sequence 1, Application US/09191233A
; GENERAL INFORMATION:
; APPLICANT: Kain, Steve
; TITLE OF INVENTION: Rapidly Degrading GFP-Fusion Proteins and Methods
; FILE REFERENCE: D6100CIP
; CURRENT APPLICATION NUMBER: US/09/191,233A
; CURRENT FILING DATE: 1998-11-13
; EARLIER APPLICATION NUMBER: US 09/062,102
; EARLIER FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 3
; SEQ ID NO 1
; LENGTH: 281
; TYPE: PRT
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: Sequence of the EGFP-MODC422-461 fusion protein.
US-09-191-233-1

Query Match 100.0%; Score 1500; DB 15; Length 281;
Best Local Similarity 100.0%; Pred. No. 1e-151;
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 1 MWSKGEELFTGVVPIVLELDGVNGHKFSVSGEGDATTYKLTIKFTCTTGKLPVPMPT 60
Db 1 MWSKGEELFTGVVPIVLELDGVNGHKFSVSGEGDATTYKLTIKFTCTTGKLPVPMPT 60
OY 61 LVTTLTLYGQCFSRPYDHNKQDFFKSNMPEGVQERTIFFKDDGNKTRAEVKEGDTL 120
Db 61 LVTTLTLYGQCFSRPYDHNKQDFFKSNMPEGVQERTIFFKDDGNKTRAEVKEGDTL 120
OY 121 VNRLEIKGIDFKEDGNILGHKLEYNVNSHNYIMADKQNGIKVNFKIRHNIEDGSVOLA 180
Db 121 VNRLEIKGIDFKEDGNILGHKLEYNVNSHNYIMADKQNGIKVNFKIRHNIEDGSVOLA 180
OY 181 DHYQONTPIGDGPVLLPDNHYLSTOSALS KDPNEKRDMVLLLEFVTAAGITLGMDELYKK 240
Db 181 DHYQONTPIGDGPVLLPDNHYLSTOSALS KDPNEKRDMVLLLEFVTAAGITLGMDELYKK 240
OY 241 LSHGPPPEVEEODDGTLPMSCAQESGMDRHPAACAASARINV 281
Db 241 LSHGPPPEVEEODDGTLPMSCAQESGMDRHPAACAASARINV 281
```

RESULT 3
US-09-191-233-1

; Sequence 1, Application US/09191233B

; GENERAL INFORMATION:
; APPLICANT: Kain, Steve
; TITLE OF INVENTION: Rapidly Degrading GFP-Fusion Proteins and Methods
; FILE REFERENCE: D6100CIP
; CURRENT APPLICATION NUMBER: US/09/191,233B
; CURRENT FILING DATE: 1998-11-13
; EARLIER APPLICATION NUMBER: US 09/062,102
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 1
; LENGTH: 281
; TYPE: PRT
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: Sequence of the EGFP-MODC422-461 fusion protein.
US-09-191-233-1

Query Match 100.0%; Score 1500; DB 15; Length 281;
Best Local Similarity 100.0%; Pred. No. 1e-151;
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 1 MWSKGEELFTGVVPIVLELDGVNGHKFSVSGEGDATTYKLTIKFTCTTGKLPVPMPT 60
Db 1 MWSKGEELFTGVVPIVLELDGVNGHKFSVSGEGDATTYKLTIKFTCTTGKLPVPMPT 60
OY 61 LVTTLTLYGQCFSRPYDHNKQDFFKSNMPEGVQERTIFFKDDGNKTRAEVKEGDTL 120
Db 61 LVTTLTLYGQCFSRPYDHNKQDFFKSNMPEGVQERTIFFKDDGNKTRAEVKEGDTL 120
OY 121 VNRLEIKGIDFKEDGNILGHKLEYNVNSHNYIMADKQNGIKVNFKIRHNIEDGSVOLA 180
Db 121 VNRLEIKGIDFKEDGNILGHKLEYNVNSHNYIMADKQNGIKVNFKIRHNIEDGSVOLA 180
OY 181 DHYQONTPIGDGPVLLPDNHYLSTOSALS KDPNEKRDMVLLLEFVTAAGITLGMDELYKK 240
Db 181 DHYQONTPIGDGPVLLPDNHYLSTOSALS KDPNEKRDMVLLLEFVTAAGITLGMDELYKK 240
OY 241 LSHGPPPEVEEODDGTLPMSCAQESGMDRHPAACAASARINV 281
Db 241 LSHGPPPEVEEODDGTLPMSCAQESGMDRHPAACAASARINV 281
```

RESULT 4
US-09-365-089-1

; Sequence 1, Application US/09365089
; GENERAL INFORMATION:
; APPLICANT: Kain, Steve
; TITLE OF INVENTION: Rapidly Degrading GFP-Fusion Proteins and Methods
; FILE REFERENCE: D6100CIP/D1
; CURRENT APPLICATION NUMBER: US/09/365,089
; CURRENT FILING DATE: 1999-07-30
; EARLIER APPLICATION NUMBER: US 09/191,233
; EARLIER FILING DATE: 1998-11-13
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 1
; LENGTH: 281
; TYPE: PRT
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: Sequence of the EGFP-MODC422-461 fusion protein.
US-09-365-089-1

Query Match 100.0%; Score 1500; DB 17; Length 281;
Best Local Similarity 100.0%; Pred. No. 1e-151;
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 1 MWSKGEELFTGVVPIVLELDGVNGHKFSVSGEGDATTYKLTIKFTCTTGKLPVPMPT 60
Db 1 MWSKGEELFTGVVPIVLELDGVNGHKFSVSGEGDATTYKLTIKFTCTTGKLPVPMPT 60
```


Db 1 MWSKGEELFTGVVPIILVELDGVNKHKFSVSGEGDATTGKLTLLKFICTTGKLPVMPPT 60
QY 61 LVTTLLTYGVQCFSRYPDHNKQHDFFKSAMPEGVQERTIFFKDDGNKTRAEVKKEGDTL 120
Db 61 LVTTLLTYGVQCFSRYPDHNKQHDFFKSAMPEGVQERTIFFKDDGNKTRAEVKKEGDTL 120
QY 121 VNRLEKGIIDFKEDGNILGHKLEYNNSHNVTYIMADKONGIKVNFKIRHNIEDSGVOIA 180
Db 121 VNRLEKGIIDFKEDGNILGHKLEYNNSHNVTYIMADKONGIKVNFKIRHNIEDSGVOIA 180
QY 181 DHYQONTPIGDGPVLLPDNHYISTQSALSKDPNEKRDHVLLEFVTAAGITLGMDELYKK 240
Db 181 DHYQONTPIGDGPVLLPDNHYISTQSALSKDPNEKRDHVLLEFVTAAGITLGMDELYKK 240
QY 241 LSHGFPPEVEEODDGLTPMSCAQESGMDRHPAACASARINV 281
Db 241 LSHGFPPEVEEODDGLTPMSCAQESGMDRHPAACASARINV 281

RESULT 5
US-09-931-232-1
; Sequence 1, Application US/09931232
; GENERAL INFORMATION:
; APPLICANT: Li, Xiangiang
; APPLICANT: Kain, Steve
; TITLE OF INVENTION: Rapidly Degrading GFP-Fusion Proteins
; TITLE OF INVENTION: Determining Ligand/Target Affinity
; FILE REFERENCE: C10075CON
; CURRENT APPLICATION NUMBER: US/09/931,232
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/364,946
; PRIOR FILING DATE: 1999-07-30
; PRIOR APPLICATION NUMBER: 09/191,233
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: 09/062,102
; PRIOR FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: 60/060,855
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 281
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Sequence of the EGFP-MODC422_461 fusion
; OTHER INFORMATION: protein.
US-09-931-232-1

Query Match 100.0%; Score 1500; DB 24; Length 281;
Best Local Similarity 100.0%; Pred. No. 1e-151;
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MWSKGEELFTGVVPIILVELDGVNKHKFSVSGEGDATTGKLTLLKFICTTGKLPVMPPT 60
Db 1 MWSKGEELFTGVVPIILVELDGVNKHKFSVSGEGDATTGKLTLLKFICTTGKLPVMPPT 60
QY 61 LVTTLLTYGVQCFSRYPDHNKQHDFFKSAMPEGVQERTIFFKDDGNKTRAEVKKEGDTL 120
Db 61 LVTTLLTYGVQCFSRYPDHNKQHDFFKSAMPEGVQERTIFFKDDGNKTRAEVKKEGDTL 120
QY 121 VNRLEKGIIDFKEDGNILGHKLEYNNSHNVTYIMADKONGIKVNFKIRHNIEDSGVOIA 180
Db 121 VNRLEKGIIDFKEDGNILGHKLEYNNSHNVTYIMADKONGIKVNFKIRHNIEDSGVOIA 180
QY 181 DHYQONTPIGDGPVLLPDNHYISTQSALSKDPNEKRDHVLLEFVTAAGITLGMDELYKK 240
Db 181 DHYQONTPIGDGPVLLPDNHYISTQSALSKDPNEKRDHVLLEFVTAAGITLGMDELYKK 240
QY 241 LSHGFPPEVEEODDGLTPMSCAQESGMDRHPAACASARINV 281
Db 241 LSHGFPPEVEEODDGLTPMSCAQESGMDRHPAACASARINV 281

RESULT 6
US-09-270-427A-37
; Sequence 37, Application US/09270427A
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Holzman, Thomas F.
; APPLICANT: Harlan, John Eric
; APPLICANT: Egan, David A.
; APPLICANT: Buko, Alexander M.
; APPLICANT: Solomon, Larry R.
; APPLICANT: Lador, Uri S.
; APPLICANT: Tang, Qiong
; TITLE OF INVENTION: CENTRIFUGALLY-ENHANCED METHOD OF
; TITLE OF INVENTION: DETERMINING LIGAND/TARGET AFFINITY
; FILE REFERENCE: 6404.US.01
; CURRENT APPLICATION NUMBER: US/09/270,427A
; CURRENT FILING DATE: 1999-03-16
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37
; LENGTH: 288
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Protein derived from plasmid pETGFP-EF3:980-1008
US-09-270-427A-37

Query Match 86.0%; Score 1290.5; DB 16; Length 288;
Best Local Similarity 95.7%; Pred. No. 3.6e-129;
Matches 243; Conservative 4; Mismatches 6; Indels 1; Gaps 1;
QY 1 MWSKGEELFTGVVPIILVELDGVNKHKFSVSGEGDATTGKLTLLKFICTTGKLPVMPPT 60
Db 21 MWSKGEELFTGVVPIILVELDGVNKHKFSVSGEGDATTGKLTLLKFICTTGKLPVMPPT 80
QY 61 LVTTLLTYGVQCFSRYPDHNKQHDFFKSAMPEGVQERTIFFKDDGNKTRAEVKKEGDTL 120
Db 81 LVTTLLTYGVQCFSRYPDHNKQHDFFKSAMPEGVQERTIFFKDDGNKTRAEVKKEGDTL 140
QY 121 VNRLEKGIIDFKEDGNILGHKLEYNNSHNVTYIMADKONGIKVNFKIRHNIEDSGVOIA 180
Db 141 VNRLEKGIIDFKEDGNILGHKLEYNNSHNVTYIMADKONGIKVNFKIRHNIEDSGVOIA 200
QY 181 DHYQONTPIGDGPVLLPDNHYISTQSALSKDPNEKRDHVLLEFVTAAGITLGMDELYKK 240
Db 201 DHYQONTPIGDGPVLLPDNHYISTQSALSKDPNEKRDHVLLEFVTAAGITLGMDELYKS 260
QY 241 LSHGFPPEVEEODD 254
Db 261 -GQAGPRIKKED 273

RESULT 7
US-10-314-861-37
; Sequence 37, Application US/10314861
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Holzman, Thomas F.
; APPLICANT: Harlan, John Eric
; APPLICANT: Egan, David A.
; APPLICANT: Buko, Alexander M.
; APPLICANT: Solomon, Larry R.
; APPLICANT: Lador, Uri S.
; APPLICANT: Tang, Qiong
; TITLE OF INVENTION: CENTRIFUGALLY-ENHANCED METHOD OF
; TITLE OF INVENTION: DETERMINING LIGAND/TARGET AFFINITY
; FILE REFERENCE: 6404.US.01
; CURRENT APPLICATION NUMBER: US/10/314,861
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: US/09/270,427
; PRIOR FILING DATE: 1999-03-16
; NUMBER OF SEQ ID NOS: 39

SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37
; LENGTH: 288
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Protein derived from plasmid pETGFP-EF3:980-1008
US-10-314-861-37

Query Match 86.0%; Score 1290.5; DB 29; Length 288;
Best Local Similarity 95.7%; Pred. No. 3.7e-129;
Matches 243; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 1 MWSKGEELFTGVVPIIVELDDGVNGHKFSVSGEGSDATYGLTLKFTCTTGKLPVPMPT 60
DB 21 MWSKGEELFTGVVPIIVELDDGVNGHKFSVSGEGSDATYGLTLKFTCTTGKLPVPMPT 80
QY 61 LVTTLTGVQCFSRYPDHMKQHDFFKSAWPGYVQERTIFFKDDGNKKTAAVKEEGDTL 120
DB 81 LVTTLTGVQCFSRYPDHMKQHDFFKSAWPGYVQERTIFFKDDGNKKTAAVKEEGDTL 140
QY 121 VNRTELKIDPKEDGNILGHKLEYNNSHNYIMADKOKNGIKVNFIRHNIEDGSVOLA 180
DB 141 VNRTELKIDPKEDGNILGHKLEYNNSHNYIMADKOKNGIKVNFIRHNIEDGSVOLA 200
QY 181 DHYOQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDMVLEFVTAAGITLGMDELYKK 240
DB 201 DHYOQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDMVLEFVTAAGITLGMDELYKS 260
QY 241 LSHGPPPEVEEODD 254
DB 261 -GGAGPRIEKED 273

RESULT 8

US-09-270-427A-35
; Sequence 35, Application US/09270427A
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Holzman, Thomas F.
; APPLICANT: Harlan, John Eric
; APPLICANT: Egan, David A.
; APPLICANT: Buko, Alexander M.
; APPLICANT: Solomon, Larry R.
; APPLICANT: Lador, Uri S.
; APPLICANT: Tang, Qiang
; TITLE OF INVENTION: CENTRIFUGALLY-ENHANCED METHOD OF
; TITLE OF INVENTION: DETERMINING LIGAND/TARGET AFFINITY
; FILE REFERENCE: 6404.US.01
; CURRENT APPLICATION NUMBER: US/09/270,427A
; CURRENT FILING DATE: 1999-03-16
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 35
; LENGTH: 293
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Protein derived from plasmid pETGFP-EF3:980-1013
US-09-270-427A-35

Query Match 86.0%; Score 1290.5; DB 16; Length 293;
Best Local Similarity 95.7%; Pred. No. 3.7e-129;
Matches 243; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 1 MWSKGEELFTGVVPIIVELDDGVNGHKFSVSGEGSDATYGLTLKFTCTTGKLPVPMPT 60
DB 21 MWSKGEELFTGVVPIIVELDDGVNGHKFSVSGEGSDATYGLTLKFTCTTGKLPVPMPT 80
QY 61 LVTTLTGVQCFSRYPDHMKQHDFFKSAWPGYVQERTIFFKDDGNKKTAAVKEEGDTL 120
DB 81 LVTTLTGVQCFSRYPDHMKQHDFFKSAWPGYVQERTIFFKDDGNKKTAAVKEEGDTL 140

QY 121 VNRTELKIDPKEDGNILGHKLEYNNSHNYIMADKOKNGIKVNFIRHNIEDGSVOLA 180
DB 141 VNRTELKIDPKEDGNILGHKLEYNNSHNYIMADKOKNGIKVNFIRHNIEDGSVOLA 200
QY 181 DHYOQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDMVLEFVTAAGITLGMDELYKK 240
DB 201 DHYOQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDMVLEFVTAAGITLGMDELYKS 260
QY 241 LSHGPPPEVEEODD 254
DB 261 -GGAGPRIEKED 273

RESULT 9

US-10-314-861-35
; Sequence 35, Application US/10314861
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Holzman, Thomas F.
; APPLICANT: Harlan, John Eric
; APPLICANT: Egan, David A.
; APPLICANT: Buko, Alexander M.
; APPLICANT: Solomon, Larry R.
; APPLICANT: Lador, Uri S.
; APPLICANT: Tang, Qiang
; TITLE OF INVENTION: CENTRIFUGALLY-ENHANCED METHOD OF
; TITLE OF INVENTION: DETERMINING LIGAND/TARGET AFFINITY
; FILE REFERENCE: 6404.US.01
; CURRENT APPLICATION NUMBER: US/10/314,861
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: US/09/270,427
; PRIOR FILING DATE: 1999-03-16
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 35
; LENGTH: 293
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Protein derived from plasmid pETGFP-EF3:980-1013
US-10-314-861-35

Query Match 86.0%; Score 1290.5; DB 29; Length 293;
Best Local Similarity 95.7%; Pred. No. 3.7e-129;
Matches 243; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 1 MWSKGEELFTGVVPIIVELDDGVNGHKFSVSGEGSDATYGLTLKFTCTTGKLPVPMPT 60
DB 21 MWSKGEELFTGVVPIIVELDDGVNGHKFSVSGEGSDATYGLTLKFTCTTGKLPVPMPT 80
QY 61 LVTTLTGVQCFSRYPDHMKQHDFFKSAWPGYVQERTIFFKDDGNKKTAAVKEEGDTL 120
DB 81 LVTTLTGVQCFSRYPDHMKQHDFFKSAWPGYVQERTIFFKDDGNKKTAAVKEEGDTL 140
QY 121 VNRTELKIDPKEDGNILGHKLEYNNSHNYIMADKOKNGIKVNFIRHNIEDGSVOLA 180
DB 141 VNRTELKIDPKEDGNILGHKLEYNNSHNYIMADKOKNGIKVNFIRHNIEDGSVOLA 200
QY 181 DHYOQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDMVLEFVTAAGITLGMDELYKK 240
DB 201 DHYOQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDMVLEFVTAAGITLGMDELYKS 260

QY 241 LSHGPPPEVEEODD 254
DB 261 -GGAGPRIEKED 273

RESULT 10

US-09-270-427A-39
; Sequence 39, Application US/09270427A
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Holzman, Thomas F.

APPLICANT: Harlan, John Eric
APPLICANT: Egan, David A.
APPLICANT: Buko, Alexander M.
APPLICANT: Solomon, Larry R.
APPLICANT: Lador, Uri S.
APPLICANT: Tang, Qiang
TITLE OF INVENTION: CENTRIFUGALLY-ENHANCED METHOD OF
TITLE OF INVENTION: DETERMINING LIGAND/TARGET AFFINITY
FILE REFERENCE: 6404 US 01
CURRENT APPLICATION NUMBER: US/09/270,427A
CURRENT FILING DATE: 1999-03-16
NUMBER OF SEQ ID NOS: 39
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 39
LENGTH: 295
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Protein derived from plasmid pETGFP-EF3:980-1008
US-09-270-427A-39

Query Match 86.0%; Score 1290.5; DB 16; Length 295;
Best Local Similarity 95.7%; Pred. No. 3.7e-129;
Matches 243; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 1 MVSKEELFTGVPIIVLELDGVDNGHKFSVSGEGSDATYGLTKLFTCTTGKLPVWPT 60
DB 21 MVSKEELFTGVPIIVLELDGVDNGHKFSVSGEGSDATYGLTKLFTCTTGKLPVWPT 80
QY 61 LVTTLYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAIVKFEEDTL 120
DB 81 LVTTLYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAIVKFEEDTL 140
QY 121 VNRLELKGIDPEKEDGNIHGKLEYNNSHNYITMADKQKGIKYNFKIRNIEDGSVOLA 180
DB 141 VNRLELKGIDPEKEDGNIHGKLEYNNSHNYITMADKQKGIKYNFKIRNIEDGSVOLA 200
QY 181 DHYQONTPIGDGVPVLLPDNHYLSTQSALSKDPNEKRDMVLLFVTAAGITLGMDELYKK 240
DB 201 DHYQONTPIGDGVPVLLPDNHYLSTQSALSKDPNEKRDMVLLFVTAAGITLGMDELYKS 260
QY 241 LSHGFPPEVEEODD 254
DB 261 -GGAGGRIRKKED 273
RESULT 11
US-10-314-861-39
Sequence 39; Application US/10314861
GENERAL INFORMATION:
APPLICANT: Abbott Laboratories
APPLICANT: Holzman, Thomas F.
APPLICANT: Harlan, John Eric
APPLICANT: Egan, David A.
APPLICANT: Buko, Alexander M.
APPLICANT: Solomon, Larry R.
APPLICANT: Lador, Uri S.
APPLICANT: Tang, Qiang
TITLE OF INVENTION: CENTRIFUGALLY-ENHANCED METHOD OF
TITLE OF INVENTION: DETERMINING LIGAND/TARGET AFFINITY
FILE REFERENCE: 6404 US 01
CURRENT APPLICATION NUMBER: US/10/314,861
CURRENT FILING DATE: 2002-12-09
PRIOR APPLICATION NUMBER: US/09/270,427
PRIOR FILING DATE: 1999-03-16
NUMBER OF SEQ ID NOS: 39
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 39
LENGTH: 295
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Protein derived from plasmid pETGFP-EF3:980-1008

US-10-314-861-39
Query Match 86.0%; Score 1290.5; DB 29; Length 295;
Best Local Similarity 95.7%; Pred. No. 3.7e-129;
Matches 243; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 1 MVSKEELFTGVPIIVLELDGVDNGHKFSVSGEGSDATYGLTKLFTCTTGKLPVWPT 60
DB 21 MVSKEELFTGVPIIVLELDGVDNGHKFSVSGEGSDATYGLTKLFTCTTGKLPVWPT 80
QY 61 LVTTLYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAIVKFEEDTL 120
DB 81 LVTTLYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAIVKFEEDTL 140
QY 121 VNRLELKGIDPEKEDGNIHGKLEYNNSHNYITMADKQKGIKYNFKIRNIEDGSVOLA 180
DB 141 VNRLELKGIDPEKEDGNIHGKLEYNNSHNYITMADKQKGIKYNFKIRNIEDGSVOLA 200
QY 181 DHYQONTPIGDGVPVLLPDNHYLSTQSALSKDPNEKRDMVLLFVTAAGITLGMDELYKK 240
DB 201 DHYQONTPIGDGVPVLLPDNHYLSTQSALSKDPNEKRDMVLLFVTAAGITLGMDELYKS 260
QY 241 LSHGFPPEVEEODD 254
DB 261 -GGAGGRIRKKED 273
RESULT 12
US-09-270-427A-33
Sequence 33; Application US/09270427A
GENERAL INFORMATION:
APPLICANT: Abbott Laboratories
APPLICANT: Holzman, Thomas F.
APPLICANT: Harlan, John Eric
APPLICANT: Egan, David A.
APPLICANT: Buko, Alexander M.
APPLICANT: Solomon, Larry R.
APPLICANT: Lador, Uri S.
APPLICANT: Tang, Qiang
TITLE OF INVENTION: CENTRIFUGALLY-ENHANCED METHOD OF
TITLE OF INVENTION: DETERMINING LIGAND/TARGET AFFINITY
FILE REFERENCE: 6404 US 01
CURRENT APPLICATION NUMBER: US/09/270,427A
CURRENT FILING DATE: 1999-03-16
NUMBER OF SEQ ID NOS: 39
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 33
LENGTH: 299
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Protein derived from plasmid pETGFP-EF3:980-1019
US-09-270-427A-33
Query Match 86.0%; Score 1290.5; DB 16; Length 299;
Best Local Similarity 95.7%; Pred. No. 3.7e-129;
Matches 243; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 241 LSHGPEPEVEEODD 254
Db 261 -GGAGPRIEKKED 273

RESULT 13
US-10-314-861-33

Sequence 33, Application US/10314861
GENERAL INFORMATION:
APPLICANT: Abbott Laboratories
APPLICANT: Holzman, Thomas F.
APPLICANT: Harlan, John Eric
APPLICANT: Egan, David A.
APPLICANT: Buko, Alexander M.
APPLICANT: Solomon, Larry R.
APPLICANT: Lador, Uri S.
APPLICANT: Tang, Qiang

TITLE OF INVENTION: CENTRIFUGALLY-ENHANCED METHOD OF
FILE REFERENCE: 6404 US 01
CURRENT FILING DATE: 2002-12-09
PRIOR FILING DATE: 1999-03-16
NUMBER OF SEQ ID NOS: 39
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 33
LENGTH: 299
TYPE: PRP
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Protein derived from plasmid pETGFP-EF3:980-1019

US-10-314-861-33

Query Match 86.0%; Score 1290.5; DB 29; Length 299;
Best Local Similarity 95.7%; Pred. No. 3.8e-129;
Matches 243; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 1 MWSKGEELFTGVVPLIVELDDGVNKHKRSVSGEGDGYGKLTIKFTCTTGKLPVMPPT 60
Db 21 MWSKGEELFTGVVPLIVELDDGVNKHKRSVSGEGDGYGKLTIKFTCTTGKLPVMPPT 80
QY 61 LVTTLTGVQCFSRYPDHMKQDFFKSAPEGYVOERTIFFKDDGNYKTRAVEREGDTL 120
Db 81 LVTTLTGVQCFSRYPDHMKQDFFKSAPEGYVOERTIFFKDDGNYKTRAVEREGDTL 140
QY 121 VNRLEKGDPEKGNILGHKLEYNVNSHNYIMADKQKNGIKVNFIRHNIEDGSYOLA 180
Db 141 VNRLEKGDPEKGNILGHKLEYNVNSHNYIMADKQKNGIKVNFIRHNIEDGSYOLA 200
QY 181 DHYQONTPIGDGPVILPDNHYLSTOSALSCKDNEKRDHVLLEFVTAAGITLGMDELYKK 240
Db 201 DHYQONTPIGDGPVILPDNHYLSTOSALSCKDNEKRDHVLLEFVTAAGITLGMDELYKS 260
QY 241 LSHGPEPEVEEODD 254
Db 261 -GGAGPRIEKKED 273

RESULT 14
US-09-270-427A-31

Sequence 31, Application US/09270427A
GENERAL INFORMATION:
APPLICANT: Abbott Laboratories
APPLICANT: Holzman, Thomas F.
APPLICANT: Harlan, John Eric
APPLICANT: Egan, David A.
APPLICANT: Buko, Alexander M.
APPLICANT: Solomon, Larry R.
APPLICANT: Lador, Uri S.
APPLICANT: Tang, Qiang
TITLE OF INVENTION: CENTRIFUGALLY-ENHANCED METHOD OF

TITLE OF INVENTION: DETERMINING LIGAND/TARGET AFFINITY
FILE REFERENCE: 6404 US 01
CURRENT APPLICATION NUMBER: US/09/270,427A
CURRENT FILING DATE: 1999-03-16
NUMBER OF SEQ ID NOS: 39
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 31
LENGTH: 305
TYPE: PRP
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Protein derived from plasmid pETGFP-EF3:980-1025

US-09-270-427A-31

Query Match 86.0%; Score 1290.5; DB 16; Length 305;
Best Local Similarity 95.7%; Pred. No. 3.9e-129;
Matches 243; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 1 MWSKGEELFTGVVPLIVELDDGVNKHKRSVSGEGDGYGKLTIKFTCTTGKLPVMPPT 60
Db 21 MWSKGEELFTGVVPLIVELDDGVNKHKRSVSGEGDGYGKLTIKFTCTTGKLPVMPPT 80
QY 61 LVTTLTGVQCFSRYPDHMKQDFFKSAPEGYVOERTIFFKDDGNYKTRAVEREGDTL 120
Db 81 LVTTLTGVQCFSRYPDHMKQDFFKSAPEGYVOERTIFFKDDGNYKTRAVEREGDTL 140
QY 121 VNRLEKGDPEKGNILGHKLEYNVNSHNYIMADKQKNGIKVNFIRHNIEDGSYOLA 180
Db 141 VNRLEKGDPEKGNILGHKLEYNVNSHNYIMADKQKNGIKVNFIRHNIEDGSYOLA 200
QY 181 DHYQONTPIGDGPVILPDNHYLSTOSALSCKDNEKRDHVLLEFVTAAGITLGMDELYKK 240
Db 201 DHYQONTPIGDGPVILPDNHYLSTOSALSCKDNEKRDHVLLEFVTAAGITLGMDELYKS 260
QY 241 LSHGPEPEVEEODD 254
Db 261 -GGAGPRIEKKED 273

RESULT 15
US-10-314-861-31

Sequence 31, Application US/10314861
GENERAL INFORMATION:
APPLICANT: Abbott Laboratories
APPLICANT: Holzman, Thomas F.
APPLICANT: Harlan, John Eric
APPLICANT: Egan, David A.
APPLICANT: Buko, Alexander M.
APPLICANT: Solomon, Larry R.
APPLICANT: Lador, Uri S.
APPLICANT: Tang, Qiang
TITLE OF INVENTION: CENTRIFUGALLY-ENHANCED METHOD OF
FILE REFERENCE: 6404 US 01
CURRENT FILING DATE: 2002-12-09
PRIOR FILING DATE: 1999-03-16
NUMBER OF SEQ ID NOS: 39
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 31
LENGTH: 305
TYPE: PRP
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Protein derived from plasmid pETGFP-EF3:980-1025

US-10-314-861-31

Query Match 86.0%; Score 1290.5; DB 29; Length 305;
Best Local Similarity 95.7%; Pred. No. 3.9e-129;
Matches 243; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 1 MWSKGEELFTGVVPLIVELDDGVNKHKRSVSGEGDGYGKLTIKFTCTTGKLPVMPPT 60

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM protein - protein search, using sw model

Run on: May 14, 2004, 08:25:49 ; Search time 20 Seconds
(Without alignments)
492.121 Million cell updates/sec

Title: US-09-931-232-1

Sequence: 1 MVSKEBELFTGVPIIVELD.....AQESGMDRHPACASARINV 281

Scoring table: BLOSUM62

Searched: 236798 seqs, 35026417 residues

Total number of hits satisfying chosen parameters: 236798

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents_AA_New:*
1: /cgn2_6/prodata/1/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/prodata/1/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/prodata/1/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/prodata/1/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/prodata/1/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/prodata/1/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/prodata/1/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1274.5	85.0	255	6	US-10-129-207B-8
2	1274	84.9	239	1	PCT-US04-10982-3
3	1274	84.9	386	6	US-10-811-081-97
4	1274	84.9	567	6	US-10-488-758-4
5	1260	84.0	491	1	PCT-US04-07434-58
6	1257	83.8	1100	1	PCT-US04-07434-48
7	1257	83.8	1132	1	PCT-US04-07434-46
8	1255	83.7	239	1	PCT-US04-07434-44
9	1255	83.4	1961	1	PCT-US04-07434-60
10	1251	83.4	238	5	US-09-786-377A-8
11	1251	83.4	238	6	US-10-333-680-3
12	1250	83.3	239	1	PCT-US04-10982-4
13	1250	83.3	239	1	PCT-US04-10982-8
14	1247	83.1	238	1	PCT-US04-10982-2
15	1247	83.1	238	6	US-10-491-013-2
16	1245	83.0	238	6	US-10-433-640-11
17	1245	83.0	239	1	PCT-US04-10982-5
18	1242	82.8	238	6	US-10-676-428-35
19	1242	82.8	238	6	US-10-491-013-3
20	1241	82.7	238	6	US-10-471-857-1
21	1240	82.7	239	1	PCT-US04-10982-6
22	1238	82.5	242	6	US-10-676-428-25
23	1238	82.5	239	1	PCT-US04-10982-7
24	1234	82.2	238	6	US-10-676-428-26
25	1233	82.2	238	6	US-10-811-081-87
26	1233	82.2	387	6	US-10-811-081-87

27	1232	82.1	238	6	US-10-676-428-29	Sequence 29, Appl
28	1229	81.9	238	6	US-10-676-428-27	Sequence 27, Appl
29	1228	81.9	238	6	US-10-676-428-28	Sequence 28, Appl
30	1228	81.9	238	6	US-10-471-857-3	Sequence 3, Appl
31	1227	81.8	238	6	US-10-676-428-36	Sequence 36, Appl
32	1227	81.8	238	6	US-10-471-857-7	Sequence 7, Appl
33	1224	81.6	238	6	US-10-676-428-23	Sequence 23, Appl
34	1224	81.6	238	6	US-10-471-857-5	Sequence 5, Appl
35	1223	81.5	238	6	US-10-676-428-31	Sequence 31, Appl
36	1223	81.5	238	6	US-10-471-857-11	Sequence 9, Appl
37	1219	81.3	238	6	US-10-676-428-32	Sequence 32, Appl
38	1215	81.0	238	6	US-10-676-428-24	Sequence 30, Appl
39	1214	80.9	238	6	US-10-676-428-30	Sequence 33, Appl
40	1214	80.9	238	6	US-10-676-428-33	Sequence 11, Appl
41	1214	80.9	238	6	US-10-471-857-11	Sequence 22, Appl
42	1213	80.9	238	6	US-10-676-428-22	Sequence 37, Appl
43	1212	80.8	238	6	US-10-676-428-37	Sequence 34, Appl
44	1211	80.7	238	6	US-10-676-428-34	Sequence 13, Appl
45	1210	80.7	238	6	US-10-471-857-13	

ALIGNMENTS

RESULT 1
US-10-129-207B-8
; Sequence 8, Application US/10129207B
; GENERAL INFORMATION:
; APPLICANT: Boehr, Gerald
; APPLICANT: Esser, Dirk
; APPLICANT: Schmidt, Ulrich
; APPLICANT: ACCT Progenomics AG
; TITLE OF INVENTION: Method for Directed Packaging of Molecular Substances
; TITLE OF INVENTION: in Protein Shells
; FILE REFERENCE: 080377-000000US
; CURRENT APPLICATION NUMBER: US/10/129,207B
; PRIOR FILING DATE: 2002-10-21
; PRIOR APPLICATION NUMBER: DE 199 52 982.5
; PRIOR FILING DATE: 1999-11-03
; PRIOR APPLICATION NUMBER: WO PCT/EP00/10878
; PRIOR FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 255
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:enhanced green
; OTHER INFORMATION: fluorescent protein (GFP) with fused C-terminal
; OTHER INFORMATION: proline-rich sequence (GFP-PRP)
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (246)..(255)
; OTHER INFORMATION: C-terminal proline-rich sequence
US-10-129-207B-8

Query Match 85.0%; Score 1274.5; DB 6; Length 255;

Best Local Similarity 97.2%; Pred. No. 4.3e-88;

Matches 241; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY	1	MVSKEBELFTGVPIIVELDGVNKHFSVSGEGSDATYKTLTKICTTGKLPVWPPT	60
DB	1	MVSKEBELFTGVPIIVELDGVNKHGKFSVSGEGSDATYKTLTKICTTGKLPVWPPT	60
QY	61	LVTTLYGVQCSRYVDHKKQDFFPSAPPEGVQERTIFFDDGNYKRAVYKEGDTL	120
DB	61	LVTTLYGVQCSRYVDHKKQDFFPSAPPEGVQERTIFFDDGNYKRAVYKEGDTL	120
QY	121	VRIELKGIQDFEDGNILGHKLEYNYSNHYVIMADKQNGIKVNRKIHNIEDGSVOLA	180
DB	121	VRIELKGIQDFEDGNILGHKLEYNYSNHYVIMADKQNGIKVNRKIHNIEDGSVOLA	180

QY 181 DHYQNTPIGDPVLLPNDHNYLSTOSALSKDNEKRDMHVLLEFVTAAGITLGMDELTK 239
DB 181 DHYQNTPIGDPVLLPNDHNYLSTOSALSKDNEKRDMHVLLEFVTAAGITLGMDELTK 240
QY 240 KLSHGPP 247
DB 241 RRASGPP 248

RESULT 2

PCT-US04-10982-3
Sequence 3, Application PC/TUS0410982
GENERAL INFORMATION:
APPLICANT: Odyssey Thera, Inc.
APPLICANT: Michnick, Stephen
APPLICANT: Macdonald, Marlene
APPLICANT: Lamerdi, Jane
TITLE OF INVENTION: FRAGMENTS OF FLUORESCENT PROTEINS FOR PROTEIN-FRAGMENT COMPLEMENT
TITLE OF INVENTION: ASSAYS
FILE REFERENCE: ODD1007
CURRENT APPLICATION NUMBER: PCT/US04/10982
CURRENT FILING DATE: 2004-04-20
PRIOR APPLICATION NUMBER: US 60/461,133
PRIOR FILING DATE: 2003-04-09
NUMBER OF SEQ ID NOS: 1067
SOFTWARE: PatentIn version 3.0
SEQ ID NO 3
LENGTH: 239
TYPE: PRT
ORGANISM: Artificial
FEATURE:
OTHER INFORMATION: "EGFP", an Aequorea-derived mutated variant of wild-type green fl
PUBLICATION INFORMATION:
AUTHORS: Zhang, J. et al
TITLE: Creating new fluorescent probes for cell biology
JOURNAL: Nat. Rev. Mol. Cell Biol.
VOLUME: 3
ISSUE: 12
PAGES: 906-918
DATE: 2002-
RELEVANT RESIDUES: (1)..(239)
PCT-US04-10982-3

Query Match 84.9%; Score 1274; DB 1; Length 239;
Best Local Similarity 100.0%; Pred. No. 4.4e-88;
Matches 239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MWSKEELFTGVVPLVLELDGVNNGHKFSVSGEGSDATYGLTKLTKCTTGKLPVMPPT 60
DB 1 MWSKEELFTGVVPLVLELDGVNNGHKFSVSGEGSDATYGLTKLTKCTTGKLPVMPPT 60
QY 61 LVTLITGVQCFSRYPDHMKOHDFPKSAMPEGVQERTIFFKDDGNKYKTRAEVKEGDTL 120
DB 61 LVTLITGVQCFSRYPDHMKOHDFPKSAMPEGVQERTIFFKDDGNKYKTRAEVKEGDTL 120
QY 121 VNRLEKIDPKEDGNILGHKLLEYNNSHNYIMADQKNGIKVNFKIRHNIEDGSVOLA 180
DB 121 VNRLEKIDPKEDGNILGHKLLEYNNSHNYIMADQKNGIKVNFKIRHNIEDGSVOLA 180
QY 181 DHYQNTPIGDPVLLPNDHNYLSTOSALSKDNEKRDMHVLLEFVTAAGITLGMDELTK 239
DB 181 DHYQNTPIGDPVLLPNDHNYLSTOSALSKDNEKRDMHVLLEFVTAAGITLGMDELTK 239

RESULT 3

US-10-811-081-97
Sequence 97, Application US/10811081
GENERAL INFORMATION:
APPLICANT: El Teyar, Nabil
APPLICANT: Campbell, Robert K
APPLICANT: Kelton, Christie A
APPLICANT: He, Chaomei

TITLE OF INVENTION: Novel Glycoproteins and Methods of Use Thereof
FILE REFERENCE: 20993-003
CURRENT APPLICATION NUMBER: US/10/811,081
CURRENT FILING DATE: 2004-03-25
PRIOR APPLICATION NUMBER: US/10/457,047
PRIOR FILING DATE: 2003-06-05
PRIOR APPLICATION NUMBER: US/10/360,149
PRIOR FILING DATE: 2003-02-06
PRIOR APPLICATION NUMBER: US/09/927,876
PRIOR FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: 60/225,035
PRIOR FILING DATE: 2000-08-11
PRIOR APPLICATION NUMBER: 60/202,724
PRIOR FILING DATE: 2000-05-08
NUMBER OF SEQ ID NOS: 107
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 97
LENGTH: 386
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Fusion Protein
US-10-811-081-97

Query Match 84.9%; Score 1274; DB 6; Length 386;
Best Local Similarity 100.0%; Pred. No. 7.4e-88;
Matches 239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MWSKEELFTGVVPLVLELDGVNNGHKFSVSGEGSDATYGLTKLTKCTTGKLPVMPPT 60
DB 148 MWSKEELFTGVVPLVLELDGVNNGHKFSVSGEGSDATYGLTKLTKCTTGKLPVMPPT 207
QY 61 LVTLITGVQCFSRYPDHMKOHDFPKSAMPEGVQERTIFFKDDGNKYKTRAEVKEGDTL 120
DB 208 LVTLITGVQCFSRYPDHMKOHDFPKSAMPEGVQERTIFFKDDGNKYKTRAEVKEGDTL 267
QY 121 VNRLEKIDPKEDGNILGHKLLEYNNSHNYIMADQKNGIKVNFKIRHNIEDGSVOLA 180
DB 268 VNRLEKIDPKEDGNILGHKLLEYNNSHNYIMADQKNGIKVNFKIRHNIEDGSVOLA 327

QY 181 DHYQNTPIGDPVLLPNDHNYLSTOSALSKDNEKRDMHVLLEFVTAAGITLGMDELTK 239
DB 328 DHYQNTPIGDPVLLPNDHNYLSTOSALSKDNEKRDMHVLLEFVTAAGITLGMDELTK 386

RESULT 4
US-10-488-758-4
Sequence 4, Application US/10488758

GENERAL INFORMATION:
APPLICANT: Howard, Andrew D.
APPLICANT: Sailer, Andreas W.
APPLICANT: Pan, Jie
APPLICANT: Fong, Tung M.
APPLICANT: Marsh, Donald J.
TITLE OF INVENTION: MELANIN-CONCENTRATING HORMONE RECEPTOR
TITLE OF INVENTION: ANTAGONIST BINDING PROTEIN
FILE REFERENCE: 20782P
CURRENT APPLICATION NUMBER: US/10/488,758
CURRENT FILING DATE: 2004-03-08
PRIOR APPLICATION NUMBER: PCT/US02/29931
PRIOR FILING DATE: 2002-09-20
PRIOR APPLICATION NUMBER: 60/325,129
PRIOR FILING DATE: 2001-09-26
NUMBER OF SEQ ID NOS: 17
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 567
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: MCH-1R antagonist binding protein
US-10-488-758-4

Query Match 84.9%; Score 1274; DB 6; Length 567;
Best Local Similarity 100.0%; Pred. No. 1.1e-87;
Matches 239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVSKEELFTGVVPLVLELDGVNNGHKEFSVSGEGEGATYKLTAKFTCTTGKLPVMPPTL 60
DB 329 MVSKEELFTGVVPLVLELDGVNNGHKEFSVSGEGEGATYKLTAKFTCTTGKLPVMPPT 388
QY 61 LVTLTYGVQCFSRYPDMKQHDFFKSAPEGYQERTIPFKDGNKYTRAIVEFGDTL 120
DB 389 LVTLTYGVQCFSRYPDMKQHDFFKSAPEGYQERTIPFKDGNKYTRAIVEFGDTL 448
QY 121 VNRLEKIDPEKEDNIIHGKLEYNVNSHNYIMADKQKNGIKVNFKIRHNIEDGSYOLA 180
DB 449 VNRLEKIDPEKEDNIIHGKLEYNVNSHNYIMADKQKNGIKVNFKIRHNIEDGSYOLA 508
QY 181 DHYOQNTPIGDGPVLLPDNHYLSTQSALSKDPNKRDMVLLFEVTAAGITLGMDELYK 239
DB 509 DHYOQNTPIGDGPVLLPDNHYLSTQSALSKDPNKRDMVLLFEVTAAGITLGMDELYK 567

RESULT 5

PCT-US04-07434-58
Sequence 58, Application PC/TUS0407434
GENERAL INFORMATION:
APPLICANT: Regents of the University of California, The
APPLICANT: Leonard, Rome H.
APPLICANT: Valerie, Kichhoefer A.
APPLICANT: Sujna, Raval-Fernandes
APPLICANT: Phoebe, Stewart L.
TITLE OF INVENTION: Vault and Vault-like Carrier Molecules
FILE REFERENCE: 14399-1PCT
CURRENT APPLICATION NUMBER: PCT/US04/07434
CURRENT FILING DATE: 2004-03-16
PRIOR APPLICATION NUMBER: 60/453,800
PRIOR FILING DATE: 2003-03-10
NUMBER OF SEQ ID NOS: 143
SOFTWARE: PatentIn version 3.2
SEQ ID NO 58
LENGTH: 491
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: synthesized and Homo sapiens
PCT-US04-07434-58

Query Match 84.0%; Score 1260; DB 1; Length 491;
Best Local Similarity 89.0%; Pred. No. 1.1e-86;
Matches 242; Conservative 4; Mismatches 20; Indels 6; Gaps 2;

QY 2 VSKGEELFTGVVPLVLELDGVNNGHKEFSVSGEGEGATYKLTAKFTCTTGKLPVMPPTL 61
DB 1 MSKEELFTGVVPLVLELDGVNNGHKEFSVSGEGEGATYKLTAKFTCTTGKLPVMPPTL 60
QY 62 VTTLTYGVQCFSRYPDMKQHDFFKSAPEGYQERTIPFKDGNKYTRAIVEFGDTL 121
DB 61 VTTLTYGVQCFSRYPDMKQHDFFKSAPEGYQERTIPFKDGNKYTRAIVEFGDTL 120
QY 122 NRIELKIDPEKEDNIIHGKLEYNVNSHNYIMADKQKNGIKVNFKIRHNIEDGSYOLA 181
DB 121 NRIELKIDPEKEDNIIHGKLEYNVNSHNYIMADKQKNGIKVNFKIRHNIEDGSYOLA 180
QY 182 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNKRDMVLLFEVTAAGITLGMDELYK 241
DB 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNKRDMVLLFEVTAAGITLGMDELYK 240
QY 242 SHGFP-----PEVEKDDGTLPMS-CAQESGM 267
DB 241 NLRIPMASALPEALCSGSRTPVLCILLESV 272

RESULT 6
PCT-US04-07434-48

Sequence 48, Application PC/TUS0407434
GENERAL INFORMATION:
APPLICANT: Regents of the University of California, The
APPLICANT: Leonard, Rome H.
APPLICANT: Valerie, Kichhoefer A.
APPLICANT: Sujna, Raval-Fernandes
APPLICANT: Phoebe, Stewart L.
TITLE OF INVENTION: Vault and Vault-like Carrier Molecules
FILE REFERENCE: 14399-1PCT
CURRENT APPLICATION NUMBER: PCT/US04/07434
CURRENT FILING DATE: 2004-03-16
PRIOR APPLICATION NUMBER: 60/453,800
PRIOR FILING DATE: 2003-03-10
NUMBER OF SEQ ID NOS: 143
SOFTWARE: PatentIn version 3.2
SEQ ID NO 48
LENGTH: 1100
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: synthesized and Rattus norvegicus
PCT-US04-07434-48

Query Match 83.8%; Score 1257; DB 1; Length 1100;
Best Local Similarity 97.5%; Pred. No. 4.2e-86;
Matches 235; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 VSKGEELFTGVVPLVLELDGVNNGHKEFSVSGEGEGATYKLTAKFTCTTGKLPVMPPTL 61
DB 1 MSKEELFTGVVPLVLELDGVNNGHKEFSVSGEGEGATYKLTAKFTCTTGKLPVMPPTL 60
QY 62 VTTLTYGVQCFSRYPDMKQHDFFKSAPEGYQERTIPFKDGNKYTRAIVEFGDTL 121
DB 61 VTTLTYGVQCFSRYPDMKQHDFFKSAPEGYQERTIPFKDGNKYTRAIVEFGDTL 120
QY 122 NRIELKIDPEKEDNIIHGKLEYNVNSHNYIMADKQKNGIKVNFKIRHNIEDGSYOLA 181
DB 121 NRIELKIDPEKEDNIIHGKLEYNVNSHNYIMADKQKNGIKVNFKIRHNIEDGSYOLA 180
QY 182 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNKRDMVLLFEVTAAGITLGMDELYK 241
DB 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNKRDMVLLFEVTAAGITLGMDELYK 240
QY 242 S 242
DB 241 A 241

RESULT 7

PCT-US04-07434-46
Sequence 46, Application PC/TUS0407434
GENERAL INFORMATION:
APPLICANT: Regents of the University of California, The
APPLICANT: Leonard, Rome H.
APPLICANT: Valerie, Kichhoefer A.
APPLICANT: Sujna, Raval-Fernandes
APPLICANT: Phoebe, Stewart L.
TITLE OF INVENTION: Vault and Vault-like Carrier Molecules
FILE REFERENCE: 14399-1PCT
CURRENT APPLICATION NUMBER: PCT/US04/07434
CURRENT FILING DATE: 2004-03-16
PRIOR APPLICATION NUMBER: 60/453,800
PRIOR FILING DATE: 2003-03-10
NUMBER OF SEQ ID NOS: 143
SOFTWARE: PatentIn version 3.2
SEQ ID NO 46
LENGTH: 1132
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: synthesized and Homo sapiens
PCT-US04-07434-46

Query Match 83.8%; Score 1257; DB 1; Length 1132;
Best Local Similarity 97.5%; Pred. No. 4.3e-86;
Matches 235; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 VSKGEELFTGVVPLIVELDGVNGHKFSVSGEGEGDGYKLTIKFTCTTGKLPVWPPTL 61
DB 1 MSKGEELFTGVVPLIVELDGVNGHKFSVSGEGEGDGYKLTIKFTCTTGKLPVWPPTL 60
QY 62 VTTLTGVQCSRSYRYPDHMKOHDFKSAPEGYOERTIFFKDDGNYKTRAVEKFGDTLV 121
DB 61 VTTLTGVQCSRSYRYPDHMKOHDFKSAPEGYOERTIFFKDDGNYKTRAVEKFGDTLV 120
QY 122 NRIELKIDFKEDGNILGHKLEYNNSHNVYIMADKQNGIKVNFKIRHNIEDGSVOLAD 181
DB 121 NRIELKIDFKEDGNILGHKLEYNNSHNVYIMADKQNGIKVNFKIRHNIEDGSVOLAD 180
QY 182 HYQONTPIGDSGPVLLPDNHYLSTQSALSKDPNEKRDMVLLFEVTAAGITGMDLYK 241
DB 181 HYQONTPIGDSGPVLLPDNHYLSTQSALSKDPNEKRDMVLLFEVTAAGITGMDLYK 240
QY 242 S 242
DB 241 A 241

RESULT 8
PCT-US04-07434-44
; Sequence 44, Application PC/TUS0407434
; GENERAL INFORMATION:
; APPLICANT: Regents of the University of California, The
; APPLICANT: Leonard, Rome H.
; APPLICANT: Valerie, Kickhoefer A.
; APPLICANT: Sujna, Raval-Fernandes
; APPLICANT: Phoebe, Stewart L.
; TITLE OF INVENTION: Vault and Vault-like Carrier Molecules
; FILE REFERENCE: 14399-1PCT
; CURRENT APPLICATION NUMBER: PCT/US04/07434
; CURRENT FILING DATE: 2004-03-16
; PRIOR APPLICATION NUMBER: 60/453,800
; PRIOR FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 44
; LENGTH: 239
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: completely synthesized
PCT-US04-07434-44

Query Match 83.7%; Score 1255; DB 1; Length 239;
Best Local Similarity 98.7%; Pred. No. 1.2e-86;
Matches 235; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 VSKGEELFTGVVPLIVELDGVNGHKFSVSGEGEGDGYKLTIKFTCTTGKLPVWPPTL 61
DB 1 MSKGEELFTGVVPLIVELDGVNGHKFSVSGEGEGDGYKLTIKFTCTTGKLPVWPPTL 60
QY 62 VTTLTGVQCSRSYRYPDHMKOHDFKSAPEGYOERTIFFKDDGNYKTRAVEKFGDTLV 121
DB 61 VTTLTGVQCSRSYRYPDHMKOHDFKSAPEGYOERTIFFKDDGNYKTRAVEKFGDTLV 120
QY 122 NRIELKIDFKEDGNILGHKLEYNNSHNVYIMADKQNGIKVNFKIRHNIEDGSVOLAD 181
DB 121 NRIELKIDFKEDGNILGHKLEYNNSHNVYIMADKQNGIKVNFKIRHNIEDGSVOLAD 180
QY 182 HYQONTPIGDSGPVLLPDNHYLSTQSALSKDPNEKRDMVLLFEVTAAGITGMDLYK 239
DB 181 HYQONTPIGDSGPVLLPDNHYLSTQSALSKDPNEKRDMVLLFEVTAAGITGMDLYK 238

RESULT 9
PCT-US04-07434-60

; Sequence 60, Application PC/TUS0407434
; GENERAL INFORMATION:
; APPLICANT: Regents of the University of California, The
; APPLICANT: Leonard, Rome H.
; APPLICANT: Valerie, Kickhoefer A.
; APPLICANT: Sujna, Raval-Fernandes
; APPLICANT: Phoebe, Stewart L.
; TITLE OF INVENTION: Vault and Vault-like Carrier Molecules
; FILE REFERENCE: 14399-1PCT
; CURRENT APPLICATION NUMBER: PCT/US04/07434
; CURRENT FILING DATE: 2004-03-16
; PRIOR APPLICATION NUMBER: 60/453,800
; PRIOR FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 60
; LENGTH: 1961
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthesized and Homo sapiens
PCT-US04-07434-60

Query Match 83.7%; Score 1255; DB 1; Length 1961;
Best Local Similarity 98.7%; Pred. No. 1.1e-85;
Matches 235; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 VSKGEELFTGVVPLIVELDGVNGHKFSVSGEGEGDGYKLTIKFTCTTGKLPVWPPTL 61
DB 1 MSKGEELFTGVVPLIVELDGVNGHKFSVSGEGEGDGYKLTIKFTCTTGKLPVWPPTL 60
QY 62 VTTLTGVQCSRSYRYPDHMKOHDFKSAPEGYOERTIFFKDDGNYKTRAVEKFGDTLV 121
DB 61 VTTLTGVQCSRSYRYPDHMKOHDFKSAPEGYOERTIFFKDDGNYKTRAVEKFGDTLV 120
QY 122 NRIELKIDFKEDGNILGHKLEYNNSHNVYIMADKQNGIKVNFKIRHNIEDGSVOLAD 181
DB 121 NRIELKIDFKEDGNILGHKLEYNNSHNVYIMADKQNGIKVNFKIRHNIEDGSVOLAD 180
QY 182 HYQONTPIGDSGPVLLPDNHYLSTQSALSKDPNEKRDMVLLFEVTAAGITGMDLYK 239
DB 181 HYQONTPIGDSGPVLLPDNHYLSTQSALSKDPNEKRDMVLLFEVTAAGITGMDLYK 238

RESULT 10
US-09-786-377A-8
; Sequence 8, Application US/09786377A
; GENERAL INFORMATION:
; APPLICANT: Loma Linda University
; APPLICANT: Szalay, Aladar A.
; APPLICANT: Wang, Yubao
; APPLICANT: Gefu, Wang-Pruski
; TITLE OF INVENTION: Method for Studying Protein Interactions in Vivo
; FILE REFERENCE: 11785-3
; CURRENT APPLICATION NUMBER: US/09/786,377A
; CURRENT FILING DATE: 1999-09-02
; PRIOR APPLICATION NUMBER: US 60/135,835
; PRIOR FILING DATE: 1999-05-24
; PRIOR APPLICATION NUMBER: US 60/099,068
; PRIOR FILING DATE: 1998-09-03
; PRIOR APPLICATION NUMBER: PCT/US99/20207
; PRIOR FILING DATE: 1999-09-02
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: humanized green fluorescence protein cDNA
US-09-786-377A-8

Query Match 83.4%; Score 1251; DB 5; Length 238;

Best Local Similarity 98.3%; Pred. No. 2.3e-86;
Matches 234; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 VSKGEELFTGVVPLVLVDGVDVNGHKFSVSGEGDATTGKLTAKFTCTTGKLPVWPPTL 61
DB 1 MSKGEELFTGVVPLVLVDGVDVNGHKFSVSGEGDATTGKLTAKFTCTTGKLPVWPPTL 60
QY 62 VTTTLYGVQCFSRYPDMKQHDFFKSAMPEGYVOERTIFFKDDGNKTRAEVKFEGDTLV 121
DB 61 VTTTLYGVQCFSRYPDMKQHDFFKSAMPEGYVOERTIFFKDDGNKTRAEVKFEGDTLV 120
QY 122 NRILKGIIDFKEDGNILGHKLEYNVNSHNYIMADKQNGIKVNFKIRHNIEDGSVOGLAD 181
DB 121 NRILKGIIDFKEDGNILGHKLEYNVNSHNYIMADKQNGIKVNFKIRHNIEDGSVOGLAD 180
QY 182 HYQONTPIGDGPVLLPDNHYISTQSALSKDPNEKRDHNVLLFEFTAAAGITLGMDELK 239
DB 181 HYQONTPIGDGPVLLPDNHYISTQSALSKDPNEKRDHNVLLFEFTAAAGITLGMDELK 238

RESULT 11

US-10-333-680-3

Sequence 3, Application US/10333680
GENERAL INFORMATION:
APPLICANT: Geiger, Benjamin
APPLICANT: Kam, Zvi
APPLICANT: Zamlir, Eli
APPLICANT: Barshelsky, Alexander D.
APPLICANT: Shutman, Michael
APPLICANT: Ben-zeev, Avri
TITLE OF INVENTION: METHODS OF ISOLATING GENES ENCODING PROTEINS OF SPECIFIC
FILE REFERENCE: 25254
CURRENT APPLICATION NUMBER: US/10/333,680
CURRENT FILING DATE: 2003-10-06
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn version 3.2
SEQ ID NO 3
LENGTH: 238
TYPE: PRT
ORGANISM: Aequorea victoria
US-10-333-680-3

Query Match 83.4%; Score 1251; DB 6; Length 238;
Best Local Similarity 98.3%; Pred. No. 2.3e-86;
Matches 234; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 VSKGEELFTGVVPLVLVDGVDVNGHKFSVSGEGDATTGKLTAKFTCTTGKLPVWPPTL 61
DB 1 MSKGEELFTGVVPLVLVDGVDVNGHKFSVSGEGDATTGKLTAKFTCTTGKLPVWPPTL 60
QY 62 VTTTLYGVQCFSRYPDMKQHDFFKSAMPEGYVOERTIFFKDDGNKTRAEVKFEGDTLV 121
DB 61 VTTTLYGVQCFSRYPDMKQHDFFKSAMPEGYVOERTIFFKDDGNKTRAEVKFEGDTLV 120
QY 122 NRILKGIIDFKEDGNILGHKLEYNVNSHNYIMADKQNGIKVNFKIRHNIEDGSVOGLAD 181
DB 121 NRILKGIIDFKEDGNILGHKLEYNVNSHNYIMADKQNGIKVNFKIRHNIEDGSVOGLAD 180
QY 182 HYQONTPIGDGPVLLPDNHYISTQSALSKDPNEKRDHNVLLFEFTAAAGITLGMDELK 239
DB 181 HYQONTPIGDGPVLLPDNHYISTQSALSKDPNEKRDHNVLLFEFTAAAGITLGMDELK 238

RESULT 12

PCT-US04-10982-4

Sequence 4, Application PC/TUS0410982
GENERAL INFORMATION:
APPLICANT: Odysey Thera, Inc.
APPLICANT: Michnick, Stephen
APPLICANT: Macdonald, Marlene
APPLICANT: Lamerdin, Jane
TITLE OF INVENTION: FRAGMENTS OF FLUORESCENT PROTEINS FOR PROTEIN-FRAGMENT COMPLEMENT

TITLE OF INVENTION: ASSAYS
FILE REFERENCE: ODDY007
CURRENT APPLICATION NUMBER: PCT/US04/10982
CURRENT FILING DATE: 2004-04-20
PRIOR APPLICATION NUMBER: US 60/461,133
PRIOR FILING DATE: 2003-04-09
NUMBER OF SEQ ID NOS: 1067
SOFTWARE: PatentIn version 3.0
SEQ ID NO 4
LENGTH: 239
TYPE: PRT
ORGANISM: Artificial
FEATURE:
OTHER INFORMATION: "EYFP", enhanced yellow fluorescent protein, an Aequorea-derived
OTHER INFORMATION: mutated variant of wild-type green fluorescent protei
PUBLICATION INFORMATION:
AUTHORS: Zhang, J. et al
TITLE: Creating new fluorescent probes for cell biology
JOURNAL: Nat. Rev. Mol. Cell Biol.
VOLUME: 3
ISSUE: 3
PAGES: 906-918
DATE: 2002-
RELEVANT RESIDUES: (1)..(239)
PCT-US04-10982-4

Query Match

Best Local Similarity 97.9%; Pred. No. 2.7e-86;
Matches 234; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPLVLVDGVDVNGHKFSVSGEGDATTGKLTAKFTCTTGKLPVWPPT 60
DB 1 MSKGEELFTGVVPLVLVDGVDVNGHKFSVSGEGDATTGKLTAKFTCTTGKLPVWPPT 60
QY 61 LVTTTLYGVQCFSRYPDMKQHDFFKSAMPEGYVOERTIFFKDDGNKTRAEVKFEGDTL 120
DB 61 LVTTTLYGVQCFSRYPDMKQHDFFKSAMPEGYVOERTIFFKDDGNKTRAEVKFEGDTL 120
QY 121 VNRILKGIIDFKEDGNILGHKLEYNVNSHNYIMADKQNGIKVNFKIRHNIEDGSVOGLA 180
DB 121 VNRILKGIIDFKEDGNILGHKLEYNVNSHNYIMADKQNGIKVNFKIRHNIEDGSVOGLA 180
QY 181 DHYQONTPIGDGPVLLPDNHYISTQSALSKDPNEKRDHNVLLFEFTAAAGITLGMDELK 239
DB 181 DHYQONTPIGDGPVLLPDNHYISTQSALSKDPNEKRDHNVLLFEFTAAAGITLGMDELK 239

RESULT 13

PCT-US04-10982-8

Sequence 8, Application PC/TUS0410982
GENERAL INFORMATION:
APPLICANT: Odysey Thera, Inc.
APPLICANT: Michnick, Stephen
APPLICANT: Macdonald, Marlene
APPLICANT: Lamerdin, Jane
TITLE OF INVENTION: FRAGMENTS OF FLUORESCENT PROTEINS FOR PROTEIN-FRAGMENT COMPLEMENT
FILE REFERENCE: ODDY007
CURRENT APPLICATION NUMBER: PCT/US04/10982
CURRENT FILING DATE: 2004-04-20
PRIOR APPLICATION NUMBER: US 60/461,133
PRIOR FILING DATE: 2003-04-09
NUMBER OF SEQ ID NOS: 1067
SOFTWARE: PatentIn version 3.0
SEQ ID NO 8
LENGTH: 239
TYPE: PRT
ORGANISM: Artificial
FEATURE:
OTHER INFORMATION: "EGFP", cyan fluorescent protein, an Aequorea-derived mutated var
OTHER INFORMATION: tant of wt green fluorescent protei
PUBLICATION INFORMATION:
AUTHORS: Zhang, J. et al

TITLE: Creating new fluorescent probes for cell biology
JOURNAL: Nat. Rev. Mol. Cell Biol.
VOLUME: 3
ISSUE: 12
PAGES: 906-918
DATE: 2002-
RELEVANT RESIDUES: (1) .. (239)
PCT-US04-10982-8

Query Match 83.1%; Score 1250; DB 1; Length 239;
Best Local Similarity 97.9%; Pred. No. 2.7e-86;
Matches 233; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 MVSKEELFTGVPLIVELDGDVNGHKFSVSGEGDGYGKLTIKFTCTTGKLPVMPPTL 60
DB 1 MVSKEELFTGVPLIVELDGDVNGHKFSVSGEGDGYGKLTIKFTCTTGKLPVMPPTL 60
QY 61 LVTTLTVGVCFSRYPDHMKOHDPFKSAMPEGYOERTIFFKDDGNYKTRAVERFEGDTLV 120
DB 61 LVTTLTVGVCFSRYPDHMKOHDPFKSAMPEGYOERTIFFKDDGNYKTRAVERFEGDTLV 120
QY 121 VNRLEKIDPEKEDGNILGHKLEYNVNSHNYIMADKOKNGIKVNFKIRHNIEDGSVQLAD 180
DB 121 VNRLEKIDPEKEDGNILGHKLEYNVNSHNYIMADKOKNGIKVNFKIRHNIEDGSVQLAD 180
QY 181 DHYOQNTPIGDGPVLLPDNHYLSTOSALSKDPNEKRDMVLEFVTAAGITLGMDELTK 239
DB 181 DHYOQNTPIGDGPVLLPDNHYLSTOSALSKDPNEKRDMVLEFVTAAGITLGMDELTK 239

RESULT 14

PCT-US04-10982-2
Sequence 2, Application PC/TUS0410982
GENERAL INFORMATION:
APPLICANT: Odyssey Thera, Inc.
APPLICANT: Michalek, Stephen
APPLICANT: Macdonald, Jane
TITLE OF INVENTION: FRAGMENTS OF FLUORESCENT PROTEINS FOR PROTEIN-FRAGMENT COMPLEMENT
FILE REFERENCE: ODDY007
CURRENT APPLICATION NUMBER: PCT/US04/10982
CURRENT FILING DATE: 2004-04-20
PRIOR APPLICATION NUMBER: US 60/461,133
PRIOR FILING DATE: 2003-04-09
NUMBER OF SEQ ID NOS: 1067
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 238
TYPE: PRT
ORGANISM: Aequorea victoria
PCT-US04-10982-2

Query Match 83.1%; Score 1247; DB 1; Length 238;
Best Local Similarity 97.9%; Pred. No. 4.5e-86;
Matches 233; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 VSKGEELFTGVPLIVELDGDVNGHKFSVSGEGDGYGKLTIKFTCTTGKLPVMPPTL 61
DB 1 MVSKEELFTGVPLIVELDGDVNGHKFSVSGEGDGYGKLTIKFTCTTGKLPVMPPTL 60
QY 62 VTTTLTVGVCFSRYPDHMKOHDPFKSAMPEGYOERTIFFKDDGNYKTRAVERFEGDTLV 121
DB 61 VTTTSTVGVCFSRYPDHMKOHDPFKSAMPEGYOERTIFFKDDGNYKTRAVERFEGDTLV 120
QY 122 NRIELKIDPEKEDGNILGHKLEYNVNSHNYIMADKOKNGIKVNFKIRHNIEDGSVQLAD 181
DB 121 NRIELKIDPEKEDGNILGHKLEYNVNSHNYIMADKOKNGIKVNFKIRHNIEDGSVQLAD 180
QY 182 HYOQNTPIGDGPVLLPDNHYLSTOSALSKDPNEKRDMVLEFVTAAGITLGMDELTK 239
DB 181 HYOQNTPIGDGPVLLPDNHYLSTOSALSKDPNEKRDMVLEFVTAAGITLGMDELTK 238

RESULT 15

US-10-491-013-2
Sequence 2, Application US/10491013
GENERAL INFORMATION:
APPLICANT: AMERSHAM BIOSCIENCES UK LIMITED
APPLICANT: STUBBS, Simon
APPLICANT: JONES, Anne R.
APPLICANT: DAVIES, June
APPLICANT: RUEHLMANN, Dietrich
APPLICANT: MICHAEL, Nigel P.
TITLE OF INVENTION: FLUORESCENT PROTEINS
FILE REFERENCE: PA0163
CURRENT APPLICATION NUMBER: US/10/491,013
CURRENT FILING DATE: 2004-03-25
PRIOR APPLICATION NUMBER: GB 0123314.7
PRIOR FILING DATE: 2001-09-28
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn version 3.2
SEQ ID NO 2
LENGTH: 238
TYPE: PRT
ORGANISM: Aequorea victoria
US-10-491-013-2

Query Match 83.1%; Score 1247; DB 6; Length 238;
Best Local Similarity 97.9%; Pred. No. 4.5e-86;
Matches 233; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 VSKGEELFTGVPLIVELDGDVNGHKFSVSGEGDGYGKLTIKFTCTTGKLPVMPPTL 61
DB 1 MVSKEELFTGVPLIVELDGDVNGHKFSVSGEGDGYGKLTIKFTCTTGKLPVMPPTL 60
QY 62 VTTTLTVGVCFSRYPDHMKOHDPFKSAMPEGYOERTIFFKDDGNYKTRAVERFEGDTLV 121
DB 61 VTTTSTVGVCFSRYPDHMKOHDPFKSAMPEGYOERTIFFKDDGNYKTRAVERFEGDTLV 120
QY 122 NRIELKIDPEKEDGNILGHKLEYNVNSHNYIMADKOKNGIKVNFKIRHNIEDGSVQLAD 181
DB 121 NRIELKIDPEKEDGNILGHKLEYNVNSHNYIMADKOKNGIKVNFKIRHNIEDGSVQLAD 180
QY 182 HYOQNTPIGDGPVLLPDNHYLSTOSALSKDPNEKRDMVLEFVTAAGITLGMDELTK 239
DB 181 HYOQNTPIGDGPVLLPDNHYLSTOSALSKDPNEKRDMVLEFVTAAGITLGMDELTK 238

Search completed: May 14, 2004, 08:31:38
Job time : 21 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 14, 2004, 23:59:10 ; Search time 3699 Seconds

(without alignments)
9959.875 Million cell updates/sec

Title: US-09-931-232-2

Sequence: 1 atggtgagcaagggcgagga.....tagatcaatgcttagatgc 850

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

```
1: gb_da:*
2: gb_htg:*
3: gb_in:*
4: gb_ov:*
5: gb_ov:*
6: gb_ov:*
7: gb_ov:*
8: gb_ov:*
9: gb_ov:*
10: gb_ov:*
11: gb_ov:*
12: gb_ov:*
13: gb_ov:*
14: gb_ov:*
15: gb_ov:*
16: gb_ov:*
17: gb_ov:*
18: gb_ov:*
19: gb_ov:*
20: gb_ov:*
21: gb_ov:*
22: gb_ov:*
23: gb_ov:*
24: gb_ov:*
25: gb_ov:*
26: gb_ov:*
27: gb_ov:*
28: gb_ov:*
29: gb_ov:*
30: gb_ov:*
31: gb_ov:*
32: gb_ov:*
33: gb_ov:*
34: gb_ov:*
35: gb_ov:*
36: gb_ov:*
37: gb_ov:*
38: gb_ov:*
39: gb_ov:*
40: gb_ov:*
41: gb_ov:*
```

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	850	100.0	850	AR112746	AR112746 Sequence
2	850	100.0	850	AR173740	AR173740 Sequence
3	850	100.0	850	BD232356	BD232356 Rapidly d
4	850	100.0	4280	AX923429	AX923429 Sequence
5	845	99.4	845	AX348042	AX348042 Sequence
6	845	99.4	845	AX353909	AX353909 Sequence
7	721.2	84.8	10417	AX617021	AX617021 Sequence
8	720.6	84.8	6418	AX490700	AX490700 Sequence
9	720.6	84.8	6418	AX801202	AX801202 Sequence
10	720.6	84.8	6418	AX814838	AX814838 Sequence
11	720	84.7	12789	AX590202	AX590202 Sequence
12	719.4	84.6	7326	AX814844	AX814844 Sequence
13	719.4	84.6	9528	AX268217	AX268217 Sequence
14	719.4	84.6	9528	AX801203	AX801203 Sequence
15	719.4	84.6	9528	AX814839	AX814839 Sequence
16	719.4	84.6	9528	AX814839	AX814839 Sequence
17	719.2	84.6	5713	AX146811	AX146811 Sequence
18	719.2	84.6	5713	BD222121	BD222121 Method an
19	719	84.6	8531	AX490699	AX490699 Sequence
20	719	84.6	8531	AX663051	AX663051 Sequence
21	719	84.6	10282	AX107919	AX107919 Sequence
22	719	84.6	10282	AX107920	AX107920 Sequence
23	719	84.6	10384	AX107917	AX107917 Sequence
24	719	84.6	10384	AX107918	AX107918 Sequence
25	719	84.6	10384	AX283601	AX283601 Sequence
26	718.8	84.6	6084	12	AX575208
27	718.6	84.5	1190	6	AX455471
28	718.6	84.5	1377	6	AX230579
29	718.6	84.5	1377	6	AX536308
30	718.6	84.5	3355	12	CVU76561
31	718.6	84.5	4151	12	CVU55761
32	718.6	84.5	4729	12	U57609
33	718.6	84.5	4733	6	AX658018
34	718.6	84.5	4733	12	CVU55762
35	718.6	84.5	4737	12	U57608
36	718.6	84.5	4810	12	AF525449
37	718.6	84.5	4843	12	AF525447
38	718.6	84.5	4843	12	AF525448
39	718.6	84.5	5063	6	BD139247
40	718.6	84.5	5184	6	BD270592
41	718.6	84.5	5451	6	AX771237
42	718.6	84.5	5497	12	AX237157
43	718.6	84.5	5619	6	AX008865
44	718.6	84.5	5619	6	BD211260
45	718.6	84.5	5710	12	AF490378

ALIGNMENTS

RESULT 1
LOCUS AR112746 850 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 2 from patent US 6130313.
ACCESSION AR112746
VERSION AR112746.1 GI:14092646
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 850)
AUTHORS Li, X. and Kain, S.
TITLE Rapidly degrading GFP-fusion proteins
JOURNAL Patent: US 6130313-A 2 10-OC-2000;
FEATURES Location/Qualifiers

source 1. .850
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 100.0%; Score 850; DB 6; Length 850;
Best Local Similarity 100.0%; Pred. No. 7.7e-99;
Matches 850; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGTGAGCAAGGCGAGAGAGCTGTTCAACGGGGGTGGGCCCATCTGTCAGACTGAGAC 60
DB 1 ATGTGAGCAAGGCGAGAGAGCTGTTCAACGGGGGTGGGCCCATCTGTCAGACTGAGAC 60
QY 61 GCGCAGCTAAACGGCCACAAGTTCAAGCTGTCCGGCGAGGGCGAGGGCGATGCCACTAC 120
DB 61 GCGCAGCTAAACGGCCACAAGTTCAAGCTGTCCGGCGAGGGCGAGGGCGATGCCACTAC 120
QY 121 GGCAGAGCTGACCTGGAAGTTCACTGCAACACCGGCAAGCTGCGCCGTGCTGCGCCAC 180
DB 121 GGCAGAGCTGACCTGGAAGTTCACTGCAACACCGGCAAGCTGCGCCGTGCTGCGCCAC 180
QY 181 CTCTGACCAACCTGACCTGAGGGGTGAGGTTCAGCGGCTACCCCGACCAATGAG 240
DB 181 CTCTGACCAACCTGACCTGAGGGGTGAGGTTCAGCGGCTACCCCGACCAATGAG 240
QY 241 CAGCAGCACTTCTTCAAGTCCGCCATGCGCCGAGGCTACGTCAGAGCGGCACTCTTC 300
DB 241 CAGCAGCACTTCTTCAAGTCCGCCATGCGCCGAGGCTACGTCAGAGCGGCACTCTTC 300
QY 301 TTCAAGAGCGAGCGGAATCTACAAGACCGGCGCGAGGTTCAGAGGGGAGCAACCTTG 360
DB 301 TTCAAGAGCGAGCGGAATCTACAAGACCGGCGCGAGGTTCAGAGGGGAGCAACCTTG 360
QY 361 GTGAACCGCATCGAGCTGAAGGGCATGCACTTCAAGAGAGAGCGGCAATCTGAGGAGC 420
DB 361 GTGAACCGCATCGAGCTGAAGGGCATGCACTTCAAGAGAGAGCGGCAATCTGAGGAGC 420
QY 421 AAGCTGAGATCAACTACCAACAGCCCAACGCTATATCATGCGCAGACAGAGAAAC 480
DB 421 AAGCTGAGATCAACTACCAACAGCCCAACGCTATATCATGCGCAGACAGAGAAAC 480
QY 481 GGCATCAAGGTGAATTTCAAGATCCGCGCAAAATCGAGAGAGCGGAGGTGAGCTCCGC 540
DB 481 GGCATCAAGGTGAATTTCAAGATCCGCGCAAAATCGAGAGAGCGGAGGTGAGCTCCGC 540
QY 541 GACCACTACAGCAGAGAACACCCCATCGGCGAGCGCCCGTCTCTCCCGACCAACAC 600
DB 541 GACCACTACAGCAGAGAACACCCCATCGGCGAGCGCCCGTCTCTCCCGACCAACAC 600
QY 601 TACCTGAGCAACCCAGTCCGCCCTTGAAGAAAGACCCCAAGAGAGCGGATCAATGCTC 660
DB 601 TACCTGAGCAACCCAGTCCGCCCTTGAAGAAAGACCCCAAGAGAGCGGATCAATGCTC 660
QY 661 CTGCTGAGATTTGTCACACCGCGCGGATCACTTCGCGATGGAAGAGCTGACAAAGAG 720
DB 661 CTGCTGAGATTTGTCACACCGCGCGGATCACTTCGCGATGGAAGAGCTGACAAAGAG 720
QY 721 CTTAGCAGATGCTTCCCGCCGAGGTGAGAGAGAGATGATGAGCACTGCTGCCATGCT 780
DB 721 CTTAGCAGATGCTTCCCGCCGAGGTGAGAGAGAGATGATGAGCACTGCTGCCATGCT 780
QY 781 TGTGCCCCAGAGAGCGGAGTGAACCGTCAACCTGAGAGCTGTGCTTCTGATGATCAAT 840
DB 781 TGTGCCCCAGAGAGCGGAGTGAACCGTCAACCTGAGAGCTGTGCTTCTGATGATCAAT 840
QY 841 GTGTAGATGC 850
DB 841 GTGTAGATGC 850
```

RESULT 2
AR173740 850 bp DNA linear PAT 17-DEC-2001
LOCUS AR173740

DEFINITION Sequence 2 from patent US 6306600.
ACCESSION AR173740
VERSION AR173740.1 GI:17914060
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 850)
TITLE Kain, S. and Li, X.
JOURNAL Rapidly degrading GFP-fusion proteins and methods of use
FEATURES Patent: US 6306600-A 2 23-OCT-2001;
Location/Qualifiers
SOURCE 1. .850
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 100.0%; Score 850; DB 6; Length 850;
Best Local Similarity 100.0%; Pred. No. 7.7e-99;
Matches 850; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGTGAGCAAGGCGAGAGAGCTGTTCAACGGGGGTGGGCCCATCTGTCAGACTGAGAC 60
DB 1 ATGTGAGCAAGGCGAGAGAGCTGTTCAACGGGGGTGGGCCCATCTGTCAGACTGAGAC 60
QY 61 GCGCAGCTAAACGGCCACAAGTTCAAGCTGTCCGGCGAGGGCGAGGGCGATGCCACTAC 120
DB 61 GCGCAGCTAAACGGCCACAAGTTCAAGCTGTCCGGCGAGGGCGAGGGCGATGCCACTAC 120
QY 121 GGCAGAGCTGACCTGGAAGTTCACTGCAACACCGGCAAGCTGCGCCGTGCTGCGCCAC 180
DB 121 GGCAGAGCTGACCTGGAAGTTCACTGCAACACCGGCAAGCTGCGCCGTGCTGCGCCAC 180
QY 181 CTCTGACCAACCTGACCTGAGGGGTGAGGTTCAGCGGCTACCCCGACCAATGAG 240
DB 181 CTCTGACCAACCTGACCTGAGGGGTGAGGTTCAGCGGCTACCCCGACCAATGAG 240
QY 241 CAGCAGCACTTCTTCAAGTCCGCCATGCGCCGAGGCTACGTCAGAGCGGCACTCTTC 300
DB 241 CAGCAGCACTTCTTCAAGTCCGCCATGCGCCGAGGCTACGTCAGAGCGGCACTCTTC 300
QY 301 TTCAAGAGCGAGCGGAATCTACAAGACCGGCGCGAGGTTCAGAGGGGAGCAACCTTG 360
DB 301 TTCAAGAGCGAGCGGAATCTACAAGACCGGCGCGAGGTTCAGAGGGGAGCAACCTTG 360
QY 361 GTGAACCGCATCGAGCTGAAGGGCATGCACTTCAAGAGAGAGCGGCAATCTGAGGAGC 420
DB 361 GTGAACCGCATCGAGCTGAAGGGCATGCACTTCAAGAGAGAGCGGCAATCTGAGGAGC 420
QY 421 AAGCTGAGATCAACTACCAACAGCCCAACGCTATATCATGCGCAGACAGAGAAAC 480
DB 421 AAGCTGAGATCAACTACCAACAGCCCAACGCTATATCATGCGCAGACAGAGAAAC 480
QY 481 GGCATCAAGGTGAATTTCAAGATCCGCGCAAAATCGAGAGAGCGGAGGTGAGCTCCGC 540
DB 481 GGCATCAAGGTGAATTTCAAGATCCGCGCAAAATCGAGAGAGCGGAGGTGAGCTCCGC 540
QY 541 GACCACTACAGCAGAGAACACCCCATCGGCGAGCGCCCGTCTCTCCCGACCAACAC 600
DB 541 GACCACTACAGCAGAGAACACCCCATCGGCGAGCGCCCGTCTCTCTCCCGACCAACAC 600
QY 601 TACCTGAGCAACCCAGTCCGCCCTTGAAGAAAGACCCCAAGAGAGCGGATCAATGCTC 660
DB 601 TACCTGAGCAACCCAGTCCGCCCTTGAAGAAAGACCCCAAGAGAGCGGATCAATGCTC 660
QY 661 CTGCTGAGATTTGTCACACCGCGCGGATCACTTCGCGATGGAAGAGCTGACAAAGAG 720
DB 661 CTGCTGAGATTTGTCACACCGCGCGGATCACTTCGCGATGGAAGAGCTGACAAAGAG 720
QY 721 CTTAGCAGATGCTTCCCGCCGAGGTGAGAGAGATGATGAGCACTGCTGCCATGCT 780
DB 721 CTTAGCAGATGCTTCCCGCCGAGGTGAGAGAGATGATGAGCACTGCTGCCATGCT 780
```


QY 121 GGCAGCTGACCTGAAAGTTATCTGACACACCGGCAAGTGGCCCTGAGCCACC 180
DB 217 GGCAGCTGACCTGAAAGTTATCTGACACACCGGCAAGTGGCCCTGAGCCACC 276
QY 181 CTGCTGACACCTGACCTTACGCGGCTGAGTGTCTTACGCCCTACCCGACCATGAAG 240
DB 277 CTGCTGACACCTGACCTTACGCGGCTGAGTGTCTTACGCCCTACCCGACCATGAAG 336
QY 241 CAGACGACCTTCTCAAGTCCGCGATGCGGAAAGGCTACGCTCAGGAGCGGACCATGTC 300
DB 337 CAGACGACCTTCTCAAGTCCGCGATGCGGAAAGGCTACGCTCAGGAGCGGACCATGTC 396
QY 301 TTCAAGACGACGCGCACTACAGACCGCGCGAGGTGAAGTTGAGGCGACACCTG 360
DB 397 TTCAAGACGACGCGCACTACAGACCGCGCGAGGTGAAGTTGAGGCGACACCTG 456
QY 361 GTGAAGCGCATCGAGCTGAAAGGCGATGACTTCAAGAGGACGCGCACTCTGGGCGAC 420
DB 457 GTGAAGCGCATCGAGCTGAAAGGCGATGACTTCAAGAGGACGCGCACTCTGGGCGAC 516
QY 421 AAGCTGAGATCACTCAACAGCGCACAGGCTATATCATGAGCGGACAGAGAAAGAC 480
DB 517 AAGCTGAGATCACTCAACAGCGCACAGGCTATATCATGAGCGGACAGAGAAAGAC 576
QY 577 GGCATCAAGGTGAACTTCAAGATCCGCGCAACATCGAGGACGCGAGCTGACCTGCC 636
DB 541 GGCATCAAGGTGAACTTCAAGATCCGCGCAACATCGAGGACGCGAGCTGACCTGCC 600
QY 637 GACCACTACCGACGCAACACCCCATCGGCGACGCGCGCTGCTGCGCCGACCAACAC 696
DB 601 TACCTGAGCAACCGCATCGGCTTGAAGCAAGACCCCAAGAGAGCGGATCATGATGTC 660
QY 697 TACCTGAGCAACCGCATCGGCTTGAAGCAAGACCCCAAGAGAGCGGATCATGATGTC 756
DB 757 CTGCTGAGATGTCGACCGCGCGGAGTCACTCTCGGATGAGAGGCTGACAAAG 816
QY 721 CTGAGCAATGCTTCCGCGGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
DB 817 CTGAGCAATGCTTCCGCGGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 876
QY 781 TGTGCCAG 840
DB 877 TGTGCCAG 936
QY 841 GTGTAGATGC 850
DB 937 GTGTAGATGC 946

RESULT 5
AX348042 845 bp DNA linear PAT 06-FEB-2002
LOCUS
DEFINITION Sequence 21 from Patent EP1172375.
ACCESSION AX348042
VERSION AX348042.1 GI:18614152
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Odenthal, M. and Jung, D.
TITLE Gene expression, genome alteration and reporter gene expression in myofibroblasts and myofibroblast-like cells by using the regulatory regions within the alpha smooth muscle alpha-actin gene
JOURNAL Odenthal, Margarete (DE)
FEATURES
source 1..845
location/Qualifiers
organism="synthetic construct"

ORIGIN
Query Match 99.4%; Score 845; DB 6; Length 845;
Best Local Similarity 100.0%; Pred. No. 3.3e-98;
Matches 845; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="EGFP"

QY 1 ATGTGAGCAAGGCGAGAGAGTCTTACCGGGGTGTGTCCTGTCGAGCTGAGAC 60
DB 1 ATGTGAGCAAGGCGAGAGAGTCTTACCGGGGTGTGTCCTGTCGAGCTGAGAC 60
QY 61 GCGAGCTAAACGCGCAACAGTTCAAGCTGTCCGCGAGAGAGAGAGAGAGAGAGAG 120
DB 61 GCGAGCTAAACGCGCAACAGTTCAAGCTGTCCGCGAGAGAGAGAGAGAGAGAGAG 120
QY 121 GGCAGCTGACCTGAAAGTTATCTGACACACCGGCAAGTGGCCCTGAGCCACC 180
DB 121 GGCAGCTGACCTGAAAGTTATCTGACACACCGGCAAGTGGCCCTGAGCCACC 180
QY 181 CTGCTGACACCTGACCTTACGCGGCTGAGTGTCTTACGCCCTACCCGACCATGAAG 240
DB 181 CTGCTGACACCTGACCTTACGCGGCTGAGTGTCTTACGCCCTACCCGACCATGAAG 240
QY 241 CAGACGACCTTCTCAAGTCCGCGATGCGGAAAGGCTATATCATGAGCGGACAGAGAAAG 300
DB 241 CAGACGACCTTCTCAAGTCCGCGATGCGGAAAGGCTATATCATGAGCGGACAGAGAAAG 300
QY 301 TTCAAGACGACGCGCACTACAGACCGCGCGAGGTGAAGTTGAGGCGACACCTG 360
DB 301 TTCAAGACGACGCGCACTACAGACCGCGCGAGGTGAAGTTGAGGCGACACCTG 360
QY 361 GTGAAGCGCATCGAGCTGAAAGGCGATGACTTCAAGAGGACGCGCACTCTGGGCGAC 420
DB 361 GTGAAGCGCATCGAGCTGAAAGGCGATGACTTCAAGAGGACGCGCACTCTGGGCGAC 420
QY 421 AAGCTGAGATCACTCAACAGCGCACAGGCTATATCATGAGCGGACAGAGAAAGAC 480
DB 421 AAGCTGAGATCACTCAACAGCGCACAGGCTATATCATGAGCGGACAGAGAAAGAC 480
QY 481 GGCATCAAGGTGAACTTCAAGATCCGCGCAACATCGAGGACGCGAGCTGACCTGCC 540
DB 481 GGCATCAAGGTGAACTTCAAGATCCGCGCAACATCGAGGACGCGAGCTGACCTGCC 540
QY 541 GACCACTACCGACGCAACACCCCATCGGCGACGCGCGCTGCTGCGCCGACCAACAC 600
DB 541 GACCACTACCGACGCAACACCCCATCGGCGACGCGCGCTGCTGCGCCGACCAACAC 600
QY 601 TACCTGAGCAACCGCATCGGCTTGAAGCAAGACCCCAAGAGAGCGGATCATGATGTC 660
DB 601 TACCTGAGCAACCGCATCGGCTTGAAGCAAGACCCCAAGAGAGCGGATCATGATGTC 660
QY 661 CTGCTGAGATGTCGACCGCGCGGAGTCACTCTCGGATGAGAGGCTGACAAAG 720
DB 661 CTGCTGAGATGTCGACCGCGCGGAGTCACTCTCGGATGAGAGGCTGACAAAG 720
QY 721 CTGAGCAATGCTTCCGCGGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
DB 721 CTGAGCAATGCTTCCGCGGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
QY 781 TGTGCCAG 840
DB 781 TGTGCCAG 840
QY 841 GTGTAGATGC 845
DB 841 GTGTAGATGC 845

RESULT 6
AX353909 845 bp DNA linear PAT 06-FEB-2002
LOCUS

QY	48	GGCATCAAGGGAAGCTTCAAGATCCGGCAACAATCGAGGACGGCAGCTGC	540
Db	1223	GGCATCAAGGGAAGCTTCAAGATCCGGCAACAATCGAGGACGGCAGCTGC	1282
QY	541	GACCACCTACGACGAAACATCCCATCGGCGAGGCCCCGTGCTTGCCCGAACATC	600
Db	1283	GACCACCTACGACGAAACATCCCATCGGCGAGGCCCCGTGCTTGCCCGAACATC	1342
QY	601	TACCTGAGCAACCCGATCCGCTCTAGGAAAGACCCCAACGAGAGCGGATCAATG	660
Db	1343	TACCTGAGCAACCCGATCCGCTCTAGGAAAGACCCCAACGAGAGCGGATCAATG	1402
QY	661	CTGCTGAGATTCTGACACGCGCGCGGATCACTCTCGCATGAGCGAGCTGTACA	720
Db	1403	CTGCTGAGATTCTGACACGCGCGCGGATCACTCTCGCATGAGCGAGCTGTACA	1462
QY	721	CTTAGCCATGAGCTTCCGCGCGAGGTGAG	750
Db	1463	CTGAGATGTCCTTGTCTAAGGAAGTTAAG	1492

	RESULT	8	
AX490700	LOCUS	6418 bp	DNA
	DEFINITION	Sequence 2 from Patent WO0229065,	linear
	ACCESSION	AX490700	PAT 16-AUG-2002
	VERSION	AX490700.1	GI:22323676

SOURCE
ORGANISM
synthetic construct
synthetic construct
artificial sequences.

AUTHORS
Kingsman, A. J., Mazarakis, N. D., Martin-Rendon, E., Azzouz, M. and

TITLE Retroviral vectors containing internal ribosomal entry sites
JOURNAL Patent: WO 0229065-A 2 11-APR-2002;
Oxford Biomedica (UK) Limited (GB)

FEATURES	Location/Qualifiers
SOURCE	1. .6418

```

SOURCE
1:10210
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Ventricular vector used in Examples"

```

Query Match	84.8%	Score 720.6	DB 6	Length 6418
Best Local Similarity	96.8%	Pred. No. 1.3e-82		
Matches 735	0	Mismatches 24	Indels 0	Gaps 0

QY	1	ATGCTGACCAAGGGGAGAGAGCTGTTCACCGGGGGTGTGCCATCTGTGTGAGCTGAC	60
Db	2690	ATGTGTACCAAGGGGAGAGAGCTGTTCACCGGGGGTGTGCCATCTGTGTGAGCTGAC	2748
QY	61	GGCGACGTAAACGGCCACAAGTTCCAGCGTGTCCGGCGAGGGCGAGTGCACCTTAC	120
Db	2750	GGCGACGTAAACGGCCACAAGTTCCAGCGTGTCCGGCGAGGGCGAGTGCACCTTAC	2809
QY	121	GGCAAGCTGACCCCTGAAGTTTCATCTGACACACACCGGCAAGCTGCCCGTCCGTGGCCAC	180
Db	2810	GGCAAGCTGACCCCTGAAGTTTCATCTGACACACACCGGCAAGCTGCCCGTCCGTGGCCAC	2865
QY	181	CTCGTGAACCACTTGAACCTTACGGCGGTGCAATGCTTCAAGCCGTACCCCGACCATGAA	240
Db	2870	CTCGTGAACCACTTGAACCTTACGGCGGTGCAATGCTTCAAGCCGTACCCCGACCATGAA	2925
QY	241	CAGACACGACTTCTTCAAGTCGCGCATGCCCCGAAAGGCTAAGTCAAGAGCGCACATCTTC	300
Db	2930	CAGACACGACTTCTTCAAGTCGCGCATGCCCCGAAAGGCTAAGTCAAGAGCGCACATCTTC	2985
QY	301	TTTCAAGACGACGCGAATCTACAAAGCCCCGCGCCGAGTGTGAAGTTGAGAGGCGACACCTTG	360
Db	2990	TTTCAAGACGACGCGAATCTACAAAGCCCCGCGCGAGTGTGAAGTTGAGAGGCGACACCTTG	3045

QY	361	GTGAACCGCANTCGAGCTGAAGGGGATGACTTCAAGAGGACCGGCAATCTCTGGGGCAC	420
Db	3050	GTGAACCGCATTCGAGCTGAAGGGGATGACTTCAAGAGGACCGGCAATCTCTGGGGCAC	3109
QY	421	AAGCTGGAAGTACAACTTACAAAGCCCAACGCTGTATATCATGCGCGCAACAGCAAGAAC	480
Db	3110	AAGTGGAGATACAACTTACAAAGCCCAACGCTGTATATCATGCGCGCAACAGCAAGAAC	3169
QY	481	GGCATCAAGGTGAATTGAAAGATCCGCCAACAATCGAGACCGGACGTCGAGCTGCC	540
Db	3170	GGCATCAAGGTGAATTGAAAGATCCGCCAACAATCGAGACCGGACGTCGAGCTGCC	3229
QY	541	GACCACTACAGAGAGAACACCCCATATGGGGGACGGGCCCCGTCGTCGCCGCAACAACAC	600
Db	3230	GACCACTACAGAGAGAACACCCCATATGGGGGACGGGCCCCGTCGTCGCCGCAACAACAC	3289
QY	601	TACCTGAGCACCACAGTCCGCGCTGAGCAAAAGACCCCAACGAGAACCGCGATCATGATC	660
Db	3290	TACCTGAGCACCACAGTCCGCGCTGAGCAAAAGACCCCAACGAGAACCGCGATCATGATC	3349
QY	661	CTGCTGAGATTCTGACCCGCCCGCGGATCATCTTCGCGATGACGAGCTGTACAAAGAG	720
Db	3350	CTGCTGAGATTCTGACCCGCCCGCGGATCATCTTCGCGATGACGAGCTGTACAAAGAG	3409
QY	721	CTTAGGCAATGCGCTCCCGCGGAGAGTGAAGAGAGAGAT	759
Db	3410	AGCGGCGCGACTCTAAGATGACCTCGAGGAATTCGAT	3448

RESULT 9
AX801202
PAGE 13 COM 0003

DEFINITION Sequence 11 from Patent WO03056022.

VERSION	AX801202.1	GI:37654035
KEYWORDS		
SOURCE	synthetic construct	

ORGANISM
synthetic construct
artificial sequences.

REFERENCE
AUTHORS
TITLE
JOURNAL

1
Radcliffe, P., Mltrophanous, K. and Themis, M.
Transgenic organism
Patent: WO 03056022-A 11 10-JUL-2003;
Oxford Biomedica (UK) Limited (GB)

```
source      1..6418  
            /organism="synthetic construct"  
            /mol_type="unassigned DNA"  
            /db_xref="taxon:5620"
```

ORIGIN	/note="PONY vector sequence"
Query Match	84.8%; Score 720.6; DB 6; Length 6418;
Best Local Similarity	96.8%; Pred. No. 1.3e-82;
Matches 735; Conservative	0; Mismatches 24; Indels 0; Gaps 0;
QY	1 ATGTGACCAAGGGGCGAGAGACTGTTCACCGGGGTGTGGCCATCTGTGTGAGCTGAGC 60
Db	2690 ATGTGTGACCAAGGGGCGAGAGACTGTTCACCGGGGTGTGGCCATCTGTGTGAGCTGTGAC 2745
QY	61 GCGCAGCGTAAACCGGCGACAAGTTCCGCTGTCGCCGCGAGGGCGAGGGCGATGGCCACTTAC 120
Db	2750 GCGCAGCGTAAACCGGCGACAAGTTCCGCTGTCGCCGCGAGGGCGAGGGCGATGGCCACTTAC 2805
QY	121 GGCAGAGCTGACCCCTTAAGTTCAATGCGACACCGGGCAAGCTTCCCGTGCCTTGGCCCMAC 180
Db	2810 GGCAGAGCTGACCCCTTAAGTTCAATGCGACACCGGGCAAGCTTCCCGTGCCTTGGCCCMAC 2865
QY	181 CTGTGTGACCAACCTGTACCTGACGGGCTGTGAGTGTCTTACGCGGTACCCCGACCACTATGAAG 240
Db	2870 CTGTGTGACCAACCTGTACCTGACGGGCTGTGAGTGTCTTACGCGGTACCCCGACCACTATGAAG 2925

Db 2930 CAGACGACCTTCTTCAAGTCCGACATGCCGAAAGGCTACGTCAGAGAGCGACCATCTTC 2989
Qy 301 TTCAAGAGCGAGCGCACTACAAAGACCCGCGCGAGTGAAGTTTCAGAGGCGACACCTTG 360
Db 2990 TTCAAGAGCGAGCGCACTACAAAGACCCGCGCGAGTGAAGTTTCAGAGGCGACACCTTG 3049
Qy 361 GTGAACCGCATCGAGCTGAAGAGGCGATGCACTTCAGAGAGCGAGCAACATCTGAGGCGAC 420
Db 3050 GTGAACCGCATCGAGCTGAAGAGGCGATGCACTTCAGAGAGCGAGCAACATCTGAGGCGAC 3109
Qy 421 AAGCTGAGTACAACTACACAGCCACACGCTATATCATGCGCGACAAAGCAAGAAC 480
Db 3110 AAGCTGAGTACAACTACACAGCCACACGCTATATCATGCGCGACAAAGCAAGAAC 3169
Qy 481 GGCATCAAGGTGAATCTTCAAGATCCGCGCAACATCGAGAGCGGACGCTGACTGCGC 540
Db 3170 GGCATCAAGGTGAATCTTCAAGATCCGCGCAACATCGAGAGCGGACGCTGACTGCGC 3229
Qy 541 GACCACTACAGCAAGAACACCCGATGAGGAGCGCGCGCGCTGCTGCGCGCAACAC 600
Db 3230 GACCACTACAGCAAGAACACCCGATGAGGAGCGCGCGCGCTGCTGCGCGCAACAC 3289
Qy 601 TACCTGAGCAACCCGATCGCGCTGAGCAAGAACCCCAAGAGAGCGGATCACTGCTC 660
Db 3290 TACCTGAGCAACCCGATCGCGCTGAGCAAGAACCCCAAGAGAGCGGATCACTGCTC 3349
Qy 661 CTGCTGAGTTCGTCAGCCGCGCGCGGATCACTCTCGGCAATGAGAGCTTACAAGAG 720
Db 3350 CTGCTGAGTTCGTCAGCCGCGCGCGGATCACTCTCGGCAATGAGAGCTTACAAGTAA 3409
Qy 721 CTTAGCCATGAGCTTCCGCGCGGAGTGAAGAGAGAGAT 759
Db 3410 AGCGGCGCGACCTTAGAGTGCACCTGAGAGATTCGAT 3448

RESULT 10
AX814838
LOCUS AX814838 6418 bp DNA linear PAT 05-DEC-2003
DEFINITION Sequence 12 from Patent W003064665.
ACCESSION AX814838
VERSION AX814838.1 GI:39104019
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Radcliffe, P.O., Miskin, J.E., Wilkes, F.J. and Mitrophanous, K.A.
TITLE Viral vector
JOURNALS Patent: WO 03064665-A 12 07-AUG-2003;
Oxford Biomedica (UK) Limited (GB)
FEATURES
source location/Qualifiers
1..6418
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Description of Artificial Sequence: Plasmid Sequence"

ORIGIN

Query Match 84.8%; Score 720.6; DB 6; Length 6418;
Best Local Similarity 96.8%; Pred. No. 1.3e-82;
Matches 735; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
Qy 1 ATGTGAGCAAGGGGAGAGAGGCTGTTACCGGGGTGGTGGCCATCTCTGGTGAAGTGAAC 60
Db 2690 ATGTGAGCAAGGGGAGAGAGGCTGTTACCGGGGTGGTGGCCATCTCTGGTGAAGTGAAC 2749
Qy 61 GGGCAGCTAAAGCGGCAAGATTGAGCGTGTCCGCGAGGCGAGGCGAGTGCACCTTAC 120
Db 2750 GGGCAGCTAAAGCGGCAAGATTGAGCGTGTCCGCGAGGCGAGGCGAGTGCACCTTAC 2809
Qy 121 GGGCAGCTAAAGCGGCAAGATTGAGCGTGTCCGCGAGGCGAGGCGAGTGCACCTTAC 180

Db 2810 GGCAGAGTGAACCTTGAAGTTCACTCTGACACACCGGAAAGCTGCGCGCTGCGCCAC 2869
Qy 181 CTGTGAGCAACCTTGAACCTTGAAGTTCACTCTGACACACCGGAAAGCTGCGCGCTGCGCCAC 240
Db 2870 CTGTGAGCAACCTTGAACCTTGAAGTTCACTCTGACACACCGGAAAGCTGCGCGCTGCGCCAC 2929
Qy 241 CAGACGACCTTCTTCAAGTCCGCGCATGCCGGAAGGCTACGTCAGAGAGCGGACACCATCTTC 300
Db 2930 CAGACGACCTTCTTCAAGTCCGCGCATGCCGGAAGGCTACGTCAGAGAGCGGACACCATCTTC 2989
Qy 301 TTCAAGAGCGAGCGCACTACAAAGACCCGCGCGAGTGAAGTTTCAGAGGCGACACCTTG 360
Db 2990 TTCAAGAGCGAGCGCACTACAAAGACCCGCGCGAGTGAAGTTTCAGAGGCGACACCTTG 3049
Qy 361 GTGAACCGCATCGAGCTGAAGAGGCGATGCACTTCAGAGAGCGAGCAACATCTGAGGCGAC 420
Db 3050 GTGAACCGCATCGAGCTGAAGAGGCGATGCACTTCAGAGAGCGAGCAACATCTGAGGCGAC 3109
Qy 421 AAGCTGAGTACAACTACACAGCCACACGCTATATCATGCGCGCAAGCAAGAAC 480
Db 3110 AAGCTGAGTACAACTACACAGCCACACGCTATATCATGCGCGCAAGCAAGAAC 3169
Qy 481 GGCATCAAGGTGAATCTTCAAGATCCGCGCAACATCGAGAGCGGACGCTGACTGCGC 540
Db 3170 GGCATCAAGGTGAATCTTCAAGATCCGCGCAACATCGAGAGCGGACGCTGACTGCGC 3229
Qy 541 GACCACTACAGCAAGAACACCCGATGAGGAGCGCGCGCGCTGCTGCGCGCAACAC 600
Db 3230 GACCACTACAGCAAGAACACCCGATGAGGAGCGCGCGCGCTGCTGCGCGCAACAC 3289
Qy 601 TACCTGAGCAACCCGATCGCGCTGAGCAAGAACCCCAAGAGAGCGGATCACTGCTC 660
Db 3290 TACCTGAGCAACCCGATCGCGCTGAGCAAGAACCCCAAGAGAGCGGATCACTGCTC 3349
Qy 661 CTGCTGAGTTCGTCAGCCGCGCGCGGATCACTCTCGGCAATGAGAGCTTACAAGAG 720
Db 3350 CTGCTGAGTTCGTCAGCCGCGCGCGGATCACTCTCGGCAATGAGAGCTTACAAGTAA 3409
Qy 721 CTTAGCCATGAGCTTCCGCGCGGAGTGAAGAGAGAGAT 759
Db 3410 AGCGGCGCGACCTTAGAGTGCACCTGAGAGATTCGAT 3448

RESULT 11
AX590202/c
LOCUS AX590202 12789 bp DNA linear PAT 24-JAN-2003
DEFINITION Sequence 9 from Patent W002083886.
ACCESSION AX590202
VERSION AX590202.1 GI:27901286
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Goossens, A. and Inz, D.
TITLE The use of genes encoding membrane transporter pumps to stimulate the production of secondary metabolites in biological cells
JOURNALS Patent: WO 02083886-A 9 24-OCT-2002;
Vlaams Interuniversitair Instituut voor Biotechnologie vzw. (BE)
FEATURES
source location/Qualifiers
1..12789
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="vector pK7MG2D"

ORIGIN

Query Match 84.7%; Score 720; DB 6; Length 12789;
Best Local Similarity 100.0%; Pred. No. 1.4e-82;
Matches 720; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGTGAGCAAGGGGAGAGAGCTTTCACCGGGGTGGTGGCCATCTCTGGTGAAGTGAAC 60

Best Local Similarity 97.9%; Pred. No.1.8e-82;									
Matches 729; Conservative 0; Mismatches 16; Indels 0; Gaps 0;									
QY	1	ATGCTGAGCAAGGCGGAGAGCTGTTACACCGGGGTGGTGGCCATCTCTGTGAGACTGGAC	60						
Db	3333	ATGCTGAGCAAGGCGGAGAGCTGTTACACGGGGTGGTGGCCATCTCTGTGAGACTGGAC	3392						
QY	61	GGCGACCTAAACGGCCCAAGTTCAAGCTGTCCGGACAGGAGCGAGGCGATGCAACCTAC	120						
Db	3393	GGCGACCTAAACGGCCCAAGTTCAAGCTGTGTCCGGACAGGAGCGAGGCGATGCAACCTAC	3455						
QY	121	GGCAAGCTGAACCTTGAAGTTTCATCTGTGACCAACCGCAAGCTGCCGTGGCCCAAC	180						
Db	3453	GGCAAGCTGAACCTTGAAGTTTCATCTGTGACCAACCGCAAGCTGCCGTGGCCCAAC	3512						
QY	181	CTGTGACCAACCTTGAACCTTGAACCTTGAACCTTGAACCTTGAACCTTGAACCTTGA	240						
Db	3513	CTGTGACCAACCTTGAACCTTGAACCTTGAACCTTGAACCTTGAACCTTGAACCTTGA	3572						
QY	241	CAGCAGCATCTTCTTCAAGTCCGCGATCCCGAAAGGCTACGTCAGAGCGACCATCTTC	300						
Db	3573	CAGCAGCATCTTCTTCAAGTCCGCGATCCCGAAAGGCTACGTCAGAGCGACCATCTTC	3632						
QY	301	TTTGAAGACCAACGGCAACTTAAAGACCCGCGCGAGGTGAAGTTTGAAGGCGACCTTG	360						
Db	3633	TTTGAAGACCAACGGCAACTTAAAGACCCGCGCGAGGTGAAGTTTGAAGGCGACCTTG	3692						
QY	361	GTTGAACCGCATTCGAGCTGAAGGGGATCGACTTCAAGAGAGACGGCAACATCTTGAGGAC	420						
Db	3693	GTTGAACCGCATTCGAGCTGAAGGGGATCGACTTCAAGAGAGACGGCAACATCTTGAGGAC	3752						
QY	421	AAGCTGAGTACAACTTAAACAAGCCACACGTTTATTCATGCGCGAACAAGACAAAC	480						
Db	3753	AAGCTGAGTACAACTTAAACAAGCCACACGTTTATTCATGCGCGAACAAGACAAAC	3812						
QY	481	GGCATCAAGGTGAACCTTCAAGATCCGCGCAACAATCGAAGACGGCAACGTGACCTGGC	540						
Db	3813	GGCATCAAGGTGAACCTTCAAGATCCGCGCAACAATCGAAGACGGCAACGTGACCTGGC	3872						
QY	541	GACCATCAACGACGAACAACCCCATCGAGGACGGCCCGTGTCTGTCCCGACAACAC	600						
Db	3873	GACCATCAACGACGAACAACCCCATCGAGGACGGCCCGTGTCTGTCCCGACAACAC	3932						
QY	601	TACTTGAGCACCAAGTCCGCTCTGAGCAAAAGACCCCAACGAGAGCGCGATGCATGCTC	660						
Db	3933	TACTTGAGCACCAAGTCCGCTCTGAGCAAAAGACCCCAACGAGAGCGCGATGCATGCTC	3992						
QY	661	CTGCTGAGTTCGTCGACCGCGCGCGGAGTCACTCTCGCATGACGAGCTGTACAAAG	720						
Db	3993	CTGCTGAGTTCGTCGACCGCGCGCGGAGTCACTCTCGCATGACGAGCTGTACAAAG	4052						
QY	721	CTTAGCCATGAGCTTCCCGCGGAGG 745							
Db	4053	AGCGGCGCGCATCTTACCTGACAG 4077							
RESULT 13									
AX814843 LOCUS AX814843 7327 bp DNA linear PAT 05-DEC-2003									
DEFINITION Sequence 17 from Patent WO03064665.									
VERSION AX814843									
KEYWORDS AX814843.1 GI:39104024									
SOURCE .									
ORGANISM synthetic construct									
REFERENCE 1 synthetic construct									
AUTHORS Radcliffe,P.O., Miskin,J.E., Wilkes,F.J. and Mitrophanous,K.A.									
TITLE Viral vector									
JOURNAL Patent: WO 0306465-A 17 07-AUG-2003;									
FEATURES Oxford Biomedica (UK) Limited (GB)									
location/Qualifiers									
SOURCE 1..7327									

RESULT 15

AX801203
LOCUS AX801203 9952 bp DNA linear PAT 13-OCT-2003
DEFINITION Sequence 12 from Patent WO03056022.
ACCESSION AX801203
VERSION AX801203.1 GI:37654036
KEYWORDS

SOURCE
ORGANISM
synthetic construct
synthetic construct
artificial sequences.

REFERENCE
AUTHORS Radcliffe, P., Mitrophanous, K. and Themis, M.
TITLE Transgenic organism
JOURNAL Patent: WO 03056022-A 12 10-JUL-2003;
Oxford Biomedica (UK) Limited (GB)
FEATURES
source location/Qualifiers

1..9952
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="pONY vector sequence"

ORIGIN

Query Match 84.6%; Score 719.4; DB 6; Length 9952;

Best Local Similarity 97.9%; Pred. No. 1.7e-82;

Matches 729; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

```
QY 1 ATGTGACGACGAGGCGGAGAGCTGTTCAACGGGGTGTGCTCCATCTGTGAGCTGAGC 60
Db 5959 ATGTGACGACGAGGCGGAGAGCTGTTCAACGGGGTGTGCTCCATCTGTGAGCTGAGC 6018
QY 61 GCGCAGCTAAACGCGCACAGTTCCGCGTCCGCGAGGGCGAGGGCGATGCCACTTAC 120
Db 6019 GCGCAGCTAAACGCGCACAGTTCCGCGTCCGCGAGGGCGAGGGCGATGCCACTTAC 6078
QY 121 GCGAAGCTGACCTTGAAGTTCTATCTGCAACACCGGCAAGCTGCCCTGCGCCAC 180
Db 6079 GCGAAGCTGACCTTGAAGTTCTATCTGCAACACCGGCAAGCTGCCCTGCGCCAC 6138
QY 181 CTGCTGACCACTCTGACCTTACGCGGTGCAAGTCTTCAAGCGCTACCCGACCAATGAAG 240
Db 6139 CTGCTGACCACTCTGACCTTACGCGGTGCAAGTCTTCAAGCGCTACCCGACCAATGAAG 6198
QY 241 CAGCAGCACTTCTTCAAGTCCGCGATGCCGAGGCTTACGTCAGAGCGCACCATCTTC 300
Db 6199 CAGCAGCACTTCTTCAAGTCCGCGATGCCGAGGCTTACGTCAGAGCGCACCATCTTC 6258
QY 301 TTCAAGGACGACGGAACCTAACAAGACCGCGCGGAGTGAAGTTGAGAGGCGACACCTTG 360
Db 6259 TTCAAGGACGACGGAACCTAACAAGACCGCGCGGAGTGAAGTTGAGAGGCGACACCTTG 6318
QY 361 GTGAACCGCATGAGCTGAAGGGCATTCGATCAAGAGGACGCGCAATCTCTGGGGAC 420
Db 6319 GTGAACCGCATGAGCTGAAGGGCATTCGATCAAGAGGACGCGCAATCTCTGGGGAC 6378
QY 421 AAGCTGAGTACACTTACAACGCGCAACGCTTATATATATGAGGCGACAGCAAGAAAC 480
Db 6379 AAGCTGAGTACACTTACAACGCGCAACGCTTATATATATGAGGCGACAGCAAGAAAC 6438
QY 481 GGCATCAAGGTGAGTCAAGATCCGCGCAACATCGAGGACGCGAGGCTGCGCC 540
Db 6439 GGCATCAAGGTGAGTCAAGATCCGCGCAACATCGAGGACGCGAGGCTGCGCC 6498
QY 541 GACCACTACCAAGACACCCCATCGGCGAGCGGCCGCTGCTGCGCCGACCAACAC 600
Db 6499 GACCACTACCAAGACACCCCATCGGCGAGCGGCCGCTGCTGCGCCGACCAACAC 6558
QY 601 TACCTGAGACCAAGTCCGCGCTGAGCAAAAGACCCCAAGAGAGCGCATCAATGCTC 660
Db 6559 TACCTGAGACCAAGTCCGCGCTGAGCAAAAGACCCCAAGAGAGCGCATCAATGCTC 6618
QY 661 CTGCTGAGATTGATGACGCGCGCGGATCACTCTCGGATGAGAGGCTGTACAAGAAAG 720
```

```
Db 6619 CTGCTGAGATTGATGACGCGCGCGGATCACTCTCGGATGAGAGGCTGTACAAGAAAG 6678
QY 721 CTGAGCATGCTTCCCGCGGAGG 745
Db 6679 AGCGCGCGCATCTTACCTGCAAG 6703
```

Search completed: May 15, 2004, 04:33:26
Job time : 3704 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 14, 2004, 19:55:16 ; Search time 436 Seconds
(without alignments)
8282.038 Million cell updates/sec

Title: US-09-931-232-2
Perfect score: 850
Sequence: 1 atggtgagcaaggcagaga.....tagcatcatgttagatgc 850

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: geneseq1980s:*
2: geneseq1990s:*
3: geneseq2000s:*
4: geneseq2001as:*
5: geneseq2001bs:*
6: geneseq2002s:*
7: geneseq2003as:*
8: geneseq2003bs:*
9: geneseq2003cs:*
10: geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	850	100.0	850	3	AAZ32767 Green flu
2	850	100.0	850	3	AAZ32767 EGFP-MODC
3	850	100.0	850	6	AAZ32767 EGFP-MODC
4	850	100.0	4278	9	AAZ32767 EGFP-MODC
5	850	100.0	4862	7	AAZ32767 EGFP-MODC
6	846	99.5	4163	3	AAZ32767 EGFP-MODC
7	846	99.5	4269	3	AAZ32767 EGFP-MODC
8	846	99.5	4300	3	AAZ32767 EGFP-MODC
9	846	99.4	845	6	AB161141 Enhanced
10	846	99.4	10417	7	AAZ32767 EGFP-MODC
11	846	99.4	6418	6	AAZ32767 EGFP-MODC
12	846	99.4	6418	6	AAZ32767 EGFP-MODC
13	846	99.4	6418	6	AAZ32767 EGFP-MODC
14	846	99.4	6418	6	AAZ32767 EGFP-MODC
15	846	99.4	6418	6	AAZ32767 EGFP-MODC
16	846	99.4	6418	6	AAZ32767 EGFP-MODC
17	846	99.4	6418	6	AAZ32767 EGFP-MODC
18	846	99.4	6418	6	AAZ32767 EGFP-MODC
19	846	99.4	6418	6	AAZ32767 EGFP-MODC
20	846	99.4	6418	6	AAZ32767 EGFP-MODC
21	846	99.4	6418	6	AAZ32767 EGFP-MODC
22	846	99.4	6418	6	AAZ32767 EGFP-MODC
23	846	99.4	6418	6	AAZ32767 EGFP-MODC

24	719	84.6	10292	4	AAZ32767	AAZ32767
25	719	84.6	10384	4	AAZ32767	AAZ32767
26	719	84.6	10384	4	AAZ32767	AAZ32767
27	719	84.6	10385	5	AAZ32767	AAZ32767
28	719	84.5	761	3	AAZ32767	AAZ32767
29	718.6	84.5	1190	6	AAZ32767	AAZ32767
30	718.6	84.5	1377	4	AAZ32767	AAZ32767
31	718.6	84.5	1377	6	AAZ32767	AAZ32767
32	718.6	84.5	1377	6	AAZ32767	AAZ32767
33	718.6	84.5	3865	5	AAZ32767	AAZ32767
34	718.6	84.5	3871	4	AAZ32767	AAZ32767
35	718.6	84.5	4729	3	AAZ32767	AAZ32767
36	718.6	84.5	4952	7	AAZ32767	AAZ32767
37	718.6	84.5	5069	5	AAZ32767	AAZ32767
38	718.6	84.5	5318	4	AAZ32767	AAZ32767
39	718.6	84.5	5451	8	AAZ32767	AAZ32767
40	718.6	84.5	5619	3	AAZ32767	AAZ32767
41	718.6	84.5	5643	7	AAZ32767	AAZ32767
42	718.6	84.5	6121	3	AAZ32767	AAZ32767
43	718.6	84.5	6427	7	AAZ32767	AAZ32767
44	718.6	84.5	7015	4	AAZ32767	AAZ32767
45	718.6	84.5	7015	4	AAZ32767	AAZ32767

ALIGNMENTS

RESULT 1	AAZ32767	AAZ32767
ID	AAZ32767	standard; cDNA; 850 BP.
AC	AAZ32767	
DT	15-SEP-2003	(revised)
DT	31-JUN-2000	(first entry)
XX	XX	Green fluorescent protein/ornithine decarboxylase fusion gene.
XX	XX	Enhanced green fluorescent protein; EGFP; ornithine decarboxylase; MODC;
XX	XX	fusion; reporter; degradation domain; C-terminal; gene expression;
XX	XX	protein localization; fluorescence; short-lived; destabilized; half-life;
XX	XX	rapid degradation; EGFP-MODC; EGFP-MODC422-461;
XX	XX	rapid turnover; determination; short-term event; repetitive event;
XX	XX	sensitive; developmental process; transient; periodic; cyclic;
XX	XX	expression; protein transport; circadian rhythm; kinetic; dynamic;
XX	XX	transcription; ds.
OS	OS	Aequorea victoria.
OS	OS	Mus sp.
OS	OS	Chimeric.
XX	XX	Key
XX	XX	Location/Qualifiers
XX	XX	1. 846
XX	XX	/tag= a
XX	XX	/product= "EGFP-MODC fusion protein"
XX	XX	WO9954348-A1.
XX	XX	28-OCT-1999.
XX	XX	13-NOV-1998; 98WO-US024323.
XX	XX	17-APR-1998; 98US-00062102.
XX	XX	(CLON-) CLONTECH LAB INC.
XX	XX	L1 X, Kain S;
XX	XX	WPL; 2000-013225/01.
XX	XX	P-PSDB; AAY50142.
XX	XX	New fusion protein useful for studying cell localization, for studying
XX	XX	cell lineage and for assaying activation or deactivation of

QY		ATGATGAGCAAGGCGGAGAGCTGTTCAACCGGGGTGGTCCCATCTGTGTGAGCTGGAC	6
Db	1	ATGTGTAGCAAGGCGGAGAGCTGTTCAACCGGGGTGGTCCCATCTGTGTGAGCTGGAC	60
QY		GGCGACGTAAACGGCCACAAGTTCAAGGTGTCCGGCGAGGCGAGGCGATGCCATTAC	12
Db	61	GGCGACGTAAACGGCCACAAGTTCAAGGTGTCCGGCGAGGCGAGGCGATGCCATTAC	12
QY	121	GGGAAGCTGACCTGAAAGTTCACTGTGACCAACCGGCAAGCTGTGCCGTGGCCCTGGCCACC	18
Db	121	GGCAAGCTGACCTGAAAGTTCACTGTGACCAACCGGCAAGCTGTGCCGTGGCCCTGGCCACC	18
QY	181	CTCGTGAACAACCTGACCTTACCGCGTGCAGTGTTCAGCGCTAACCCGACCACTATGAAG	24
Db	181	CTCGTGAACAACCTGACCTTACCGCGTGCAGTGTTCAGCGCTAACCCGACCACTATGAAG	24
QY	241	CAGACGACATTTCTTCAAGTCCGCGATATGCCGAAGGCTAGTCCAGGAGGCAACATCTTTC	30
Db	241	CAGACGACATTTCTTCAAGTCCGCGATATGCCGAAGGCTAGTCCAGGAGGCAACATCTTTC	30
QY	301	TTTCAAGACACACGGCCACTTCAAGAAGCCCGCGCGAGGTGAAGTTGAGAGGCGACACCCCTG	36
Db	301	TTTCAAGACACACGGCCACTTCAAGAAGCCCGCGCGAGGTGAAGTTGAGAGGCGACACCCCTG	36
QY	361	GTTGAACCGCATCGAGCTGAAGGGCATTCGCTTCAAGAGGACCGGCAACAATCTTGGGGCAC	42
Db	361	GTTGAACCGCATCGAGCTGAAGGGCATTCGCTTCAAGAGGACCGGCAACAATCTTGGGGCAC	42
QY	421	AAGCTGAGAGTACAACTTCAACAGCCCAACAGTGTATATCATGGCCGCAACGACGAGGAAC	48
Db	421	AAGCTGAGAGTACAACTTCAACAGCCCAACAGTGTATATCATGGCCGCAACGAGGAAC	48
QY	481	GGCATCAAGGTGAATTCAGATTCGCGCAACAATCGAGGACGGCAGCGTCACTCGCC	54
Db	481	GGCATCAAGGTGAATTCAGATTCGCGCAACAATCGAGGACGGCAGCGTCACTCGCC	54
QY	541	GACCATTAACAAGCAACAATCCCATATGCGACGCGCCGCTGTGCTGCCCGAACCAAC	60
Db	541	GACCATTAACAAGCAACAATCCCATATGCGACGCGCCGCTGTGCTGCCCGAACCAAC	60
QY	601	TACCTGAGCAACCCAGTCCGCCCTGAGCAAAAGCCCAACGAGAAAGGCGATCATGATTC	66
Db	601	TACCTGAGCAACCCAGTCCGCCCTGAGCAAAAGCCCAACGAGAAAGGCGATCATGATTC	66
QY	661	CTGCTGAGATTCGTGACCGCGCGCGGAGATCATCTCGGCAATGACAGCTGTACAAGAA	72
Db	661	CTGCTGAGATTCGTGACCGCGCGCGGAGATCATCTCGGCAATGACAGCTGTACAAGAA	72
QY	721	CTTAGGCATGGCTTCCCGCGAGGTGGAAGACAGGATGATGAGCAAGCGTGTCTGTGTAAGATCAAT	78
Db	721	CTTAGGCATGGCTTCCCGCGAGGTGGAAGACAGGATGATGAGCAAGCGTGTGTCTGTGTAAGATCAAT	78
QY	781	TGTGCCCCAGAGAGACGGGATGGAACCGTCAACCTTGACGCTGTCTTCTGTGTAAGATCAAT	84
Db	781	TGTGCCCCAGAGAGACGGGATGGAACCGTCAACCTTGACGCTGTCTTCTGTGTAAGATCAAT	84
QY	841	GTTGATGATGC 850	
Db	841	GTTGATGATGC 850	
RESULT 4			
ADE35067			
ID	ADE35067 standard; DNA, 4278 BP.		
XX	ADE35067;		
XX	29-JAN-2004 (first entry)		
XX	Plasmid pd2BGF-1.		
XX	stem cell; green fluorescent protein; pd2BGF, gene; ds.		

OS	Chimeric-
XX	Aequorea victoria.
FN	W02003080816-A2.
XX	
PD	02-OCT-2003.
XX	
PF	18-MAR-2003; 2003MO-GB001111.
XX	
PR	19-MAR-2002; 2002GB-00006422.
FR	08-MAY-2002; 2002GB-00010458.
XX	
PA	(UYSH-) UNIV SHEFFIELD.
PI	Andrews P, Draper J;
XX	
DR	WPI; 2003-779256/73.
XX	
PT	Manipulating phenotype of stem cell by providing cell transfected with nucleic acid comprising promoter which confers substantial stem cell specific expression on selective marker gene(s), and proliferating cell. Disclosure; Fig 5; 40pp; English.
PS	
XX	
CC	The present sequence is that of plasmid pΔBGP-1 comprising an enhanced green fluorescent protein (GFP) reporter gene. This expression vector was used to demonstrate a method of manipulating the phenotype of a stem cell. In this method, a cell is transfected with a nucleic acid molecule that includes a promoter which comprises a motif ADE35056-ADE35060 that confers stem cell specific expression on a selectable marker (e.g. fluorescent protein or prodng activating polypeptide) gene, and conditions conducive to proliferation of the cell are provided. A cell culture system is provided which facilitates the maintenance of stem cells, particularly embryonic stem cells, in an undifferentiated state. Also provided are differentiated cells and tissues, the genome of which includes a nucleic acid construct comprising a promoter which has a stem cell specific expression pattern which controls expression of a gene the expression of which allows the selective ablation of cells which have de- differentiated to a stem cell phenotype, thereby allowing their removal from a population of differentiated cells. The stem cells are haematopoietic, neural, bone, muscle, mesenchymal, trophoblastic, epithelial, endometrial or embryonic stem cells (including embryonal carcinoma TE8A2 or NT8A2 cells), or embryonal germ cells.
CC	
SQ	Sequence 4278 BP; 1015 A; 1177 C; 1150 G; 936 T; 0 U; 0 Other;
XX	
Query Match	100.0%; Score 850; DB 9; Length 4278;
Best Local Similarity	100.0%; Pred. No.1,2e-138;
Matches	850; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 ATGTGACGAAGGGCGAGAGACTGTTCACCGGGGTGTGCCATCTTGTGAGCTGCAC 60
DB	97 ATGTGAGCAAGGGCGAGAGACTGTTCACCGGGGTGTGCCATCTTGTGAGCTGCAC 156
QY	61 GGCGACGTAAACGGCGACCAATTGACGCTGCCGCGAGGGCGAGGGCATGCCACTTAC 120
DB	157 GGCGACGTAAACGGCGACCAATTGACGCTGCCGCGAGGGCGAGGGCATGCCACTTAC 216
QY	121 GGCAAGCTGACCCTTGAAAGTTATCTGCAACAACCGGCAAGCTGCCCTGCTGCCAC 180
DB	217 GGCAAGCTGACCCTTGAAAGTTATCTGCAACAACCGGCAAGCTGCCCTGCTGCCAC 276
QY	181 CTGTGACCAACCTGACCTTAAGCGGTGAGAGCTTCACACCGCTAACCGCACCATGAG 240
DB	277 CTGTGACCAACCTGACCTTAAGCGGTGAGAGCTTCACACCGCTAACCGCACCATGAG 336
QY	241 CAGCAGCACTTCTTCAAGTCCGCATGCGCAAGGCTACGTCCAGAGCGACCACTTTTC 300
DB	337 CAGCAGCACTTCTTCAAGTCCGCATGCGCAAGGCTACGTCCAGAGCGACCACTTTTC 396
QY	301 TTCAAAGACGACGGCAATTAAGAAGCCGGCGCCAGAGGTGAAGTTGAGAGGGGACACCCCTG 360
DB	397 TTCAAAGACGACGGCAATTAAGAAGCCGGCGCCAGAGGTGAAGTTGAGAGGGGACACCCCTG 456

QY 361 GTGAACCGCATCGAGCTGAAAGGCGATCTCAAGAGGACGGCAACATCTGGGGCAC 420
XX
DB 457 GTGAACCGCATCGAGCTGAAAGGCGATCTCAAGAGGACGGCAACATCTGGGGCAC 516
XX
QY 421 AAGCTGAGTACCACTCAACAGGCAACAGTCTATATCATGTGGCCGCAAGCAAGAAAC 480
DB 517 AAGCTGAGTACCACTCAACAGGCAACAGTCTATATCATGTGGCCGCAAGCAAGAAAC 576
QY 481 GGATCAAGGTGAACCTCAAGATCCGCAACATCGAGAGCGGACGTCGAGCTGCC 540
DB 577 GGATCAAGGTGAACCTCAAGATCCGCAACATCGAGAGCGGACGTCGAGCTGCC 636
QY 541 GACCACTACCAAGCAAGAACACCCCATCGGAGCGGCGGTCGTGCTGCCGCAACACAC 600
DB 637 GACCACTACCAAGCAAGAACACCCCATCGGAGCGGCGGTCGTGCTGCCGCAACACAC 696
QY 601 TACCTGAGCAACCAAGTCCGCTGAGCAAGAACACCCCAAGCAAGAGCGGATCATGTGC 660
DB 697 TACCTGAGCAACCAAGTCCGCTGAGCAAGAACACCCCAAGCAAGAGCGGATCATGTGC 756
QY 661 CTGCTGAGTTCGTCAGCGCGCGCGGATCATCTCGGCAATGAGCAAGCTGTCAAGAG 720
DB 757 CTGCTGAGTTCGTCAGCGCGCGCGGATCATCTCGGCAATGAGCAAGCTGTCAAGAG 816
QY 721 CTGACCATGCTTCCCGCGGAGGTGAGAGAGAGATGAGCAAGCTGCCATGTCT 780
DB 817 CTGACCATGCTTCCCGCGGAGGTGAGAGAGAGATGAGCAAGCTGCCATGTCT 876
QY 781 TGTGCGCAGAGAGCGGAGTGAACCCCTGACCCCTGCTGCTTGTGTAAGATCAAT 840
DB 877 TGTGCGCAGAGAGCGGAGTGAACCCCTGACCCCTGCTGCTTGTGTAAGATCAAT 936
QY 841 GTGTAGATGC 850
DB 937 GTGTAGATGC 946

RESULT 5
ACCA4691
ID ACCA4691 standard; DNA; 4862 BP.
XX
ACCA4691;

29-MAY-2003 (first entry)

Plasmid pD2eGFP-N1 nucleotide fragment SEQ ID NO:87.

XX Chromosome-based platform; artificial chromosome; eukaryotic chromosome;
XX att site; integrase; recombinase; Acs; gene therapy; transgenic animal;
XX Platform artificial chromosome expression system; gene; ds.
XX Synthetic.

XX WO200297059-A2.

XX 05-DEC-2002.

XX 30-MAY-2002; 2002WO-US017452.

XX 30-MAY-2001; 2001US-0294758P.

XX 21-MAR-2002; 2002US-0366891P.

XX (CHRO-) CHROMOS MOLECULAR SYSTEMS INC.

XX Perkins E, Perez C, Lindenbaum M, Greene A, Leung J, Fleming E;
XX Stewart S, Sheldahl J;

XX WPI; 2003-140461/13.

XX Novel eukaryotic chromosome comprising one or many att sites which
XX permits site-directed integration in the presence of lambda-integrase,
XX useful for site-specific recombination-directed integration of DNA of

PT interest.

XX Example 2; Page 209-210; 272pp; English.

XX The present invention describes a eukaryotic chromosome (I) comprising
XX one or several att sites, where an att site is heterologous to the
XX chromosome, and permits site-directed integration in the presence of
XX lambda-integrase. Also described: (1) a platform artificial chromosome
XX expression system (Acs) (II) comprising several sites that participate
XX in recombinase catalyzed recombination; and (2) a method (M1) for
XX introducing a heterologous nucleic acid into a platform artificial
XX chromosome. (I) can be used in gene therapy. (M1) is useful for
XX introducing a heterologous nucleic acid molecule into a platform
XX artificial chromosome, preferably an Acs. (II) is useful for producing a
XX transgenic animal (e.g. a fish, insect, reptile, amphibian, arachnid, or
XX mammal) by introducing (II) by cell fusion, lipid-mediated transfection
XX by a carrier system, microinjection, microcell fusion, electroporation,
XX microprojectile bombardment or direct DNA transfer into an embryonic
XX cell, preferably a stem cell or an embryo. (II) comprises a heterologous
XX nucleic acid that encodes a therapeutic product which is useful for
XX making a library of Acs comprising random portions of a genome. ACCA4612
XX to ACCA4732 and ABP96650 to ABP96657 represent sequences used in the
XX exemplification of the present invention

XX Sequence 4862 BP; 1162 A; 1291 G; 1089 T; 0 U; 0 Other;

XX Query Match 100.0%; Score 850; DB 7; Length 4862;

XX Best Local Similarity 100.0%; Pred. No. 1,2e-138; Matches 850; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATGTGAGCAAGGCGGAGAGCTGTTACCGGGGTTGTTCCATCTGTGAGCTGAGAC 60
DB 679 AATGTGAGCAAGGCGGAGAGCTGTTACCGGGGTTGTTCCATCTGTGAGCTGAGAC 738
QY 61 GGCGAGTAAACGGCCCAAGTTGACCGTTCGGGAGAGGCGAGGCGGATCACTTAC 120
DB 739 GGCGAGTAAACGGCCCAAGTTGACCGTTCGGGAGAGGCGAGGCGGATCACTTAC 798
QY 121 GGCAAGCTACCTGTAAGTTCTGACCAACCGGCAAGCTGCGCTGCGGCCAC 180
DB 799 GGCAAGCTACCTGTAAGTTCTGACCAACCGGCAAGCTGCGCTGCGGCCAC 858
QY 181 CTCTGTAACCACTGTAAGCTGTAAGCTGTAAGCTGTAAGCTGTAAGCTGTAAG 240
DB 859 CTCTGTAACCACTGTAAGCTGTAAGCTGTAAGCTGTAAGCTGTAAGCTGTAAG 918
QY 241 CAGCAGACTTCTTAAGTCCGCGATGCGGCAAGCTGTAAGCTGTAAGCTGTAAG 300
DB 919 CAGCAGACTTCTTAAGTCCGCGATGCGGCAAGCTGTAAGCTGTAAGCTGTAAG 978
QY 301 TTCAAGAGAGAGCGGCAACTCAAGACCCGCGGAGTGAAGTTGAGGCGGCAACCTTG 360
DB 979 TTCAAGAGAGAGCGGCAACTCAAGACCCGCGGAGTGAAGTTGAGGCGGCAACCTTG 1038
QY 361 GTGAACCGCATCGAGCTGAAAGGCGATCTCAAGAGGACGGCAACATCTGGGGCAC 420
DB 1039 GTGAACCGCATCGAGCTGAAAGGCGATCTCAAGAGGACGGCAACATCTGGGGCAC 1098
QY 421 AAGCTGAGTACCACTCAACAGGCAACAGTCTATATCATGTGGCCGCAAGCAAGAAAC 480
DB 1099 AAGCTGAGTACCACTCAACAGGCAACAGTCTATATCATGTGGCCGCAAGCAAGAAAC 1158
QY 481 GGATCAAGGTGAACCTCAAGATCCGCAACATCGAGAGCGGACGTCGAGCTGCC 540
DB 1159 GGATCAAGGTGAACCTCAAGATCCGCAACATCGAGAGCGGACGTCGAGCTGCC 1218
QY 541 GACCACTACCAAGCAAGAACACCCCATCGGAGCGGCGGTCGTGCTGCCGCAACACAC 600
DB 1219 GACCACTACCAAGCAAGAACACCCCATCGGAGCGGCGGTCGTGCTGCCGCAACACAC 1278
QY 601 TACCTGAGCAACCAAGTCCGCTGAGCAAGAACACCCCAAGCAAGAGCGGATCATGTGC 660
DB 1279 TACCTGAGCAACCAAGTCCGCTGAGCAAGAACACCCCAAGCAAGAGCGGATCATGTGC 1338

QY 661 CTGCTGAGCTTGTGACCGCGCGCGGATCACTTCGCGANTGACGAGCTGTACAAGAG 720
DB 1339 CTGCTGAGAGTTGCTGACCGCGCGCGGATCACTTCGCGANTGACGAGCTGTACAAGAG 1398
QY 721 CTTAGCCAGAGCTTCCCGCGCGAGGTGAGAGGAGAGATGATGAGACCGTGGCCATGTCT 780
DB 1399 CTTAGCCAGAGCTTCCCGCGCGAGGTGAGAGGAGAGATGATGAGACCGTGGCCATGTCT 1458
QY 781 TGTGCCAGAGAGAGCGGAGTGAACCGTCAACCTGTGAGCTGTGTCTTGTGATGATCAAT 840
DB 1459 TGTGCCAGAGAGAGCGGAGTGAACCGTCAACCTGTGAGCTGTGTCTTGTGATGATCAAT 1518
QY 841 GTGTAGATGC 850
DB 1519 GTGTAGATGC 1528

RESULT 6
AAA29134
ID AAA29134 standard; DNA; 4163 BP.

AC AAA29134;
XX
DT 12-SEP-2000 (first entry)
XX
DE PNF-kappaB-d2EGFP construct.
XX
KW Vector; transcription; cis-element; nuclear factor-kappaB; NF-kappaB;
KM d2EGFP; red-shifted; destabilized; green fluorescent protein; reporter;
XX promoter; monitor; signal transduction; ss.
OS Synthetic.
XX
PN WO200034435-A2.
XX
PD 15-JUN-2000.
XX
PF 01-DEC-1999; 99WO-US028451.
XX
PR 08-DEC-1998; 98US-00206887.
XX
PA (CLON-) CLONTECH LAB INC.
XX
PI Li X, Zhao X, Jiang X, Fang Y, Duong T, Kain S;
XX
DR WPI; 2000-423400/36.
XX
PT Cis element-reporter construct for measuring transcription, comprises a
PT reporter gene, promoter and a cis element such as nuclear factor-kappaB,
PT heat responsive element, Myc, p53 or alkaline phosphatase 1.
XX
PS Claim 6; Page 34-37; 42pp; English.
XX

CC A new construct for measuring transcription comprises a cis-element from
CC a DNA sequence such as that for nuclear factor-kappaB (NF-kappaB), heat
CC responsive element (HRE), Myc, p53, nuclear factor of activated T-cells
CC (NFAT), activator protein 1 (AP-1), serum responsive element (SRE),
CC glucocorticoid responsive element (GRE) or cyclic adenosine monophosphate
CC (cAMP) responsive element, and a reporter gene and a promoter. Reporter
CC genes are chosen from secreted alkaline phosphatase (SEAP), destabilized
CC green fluorescent protein (d2EGFP) and luciferase (Luc). Promoter are
CC preferably chosen from the thymidine kinase and gonadotropin-alpha gene
CC promoters. The constructs are useful for measuring transcription and for
CC monitoring the activation of a transcription factor. They are also used
CC to determine whether a gene is involved in signal transduction pathway,
CC where the cis-element is a component of a DNA sequence which affects the
CC expression of a test gene
XX

Sequence 4163 BP; 1033 A; 1097 C; 1039 G; 994 T; 0 U; 0 Other;

Query Match 99.5%; Score 846; DB 3; Length 4163;
Best Local Similarity 100.0%; Pred. No. 5,9e-138;

Matches 846; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTGACAGAGGCGAGAGAGCTTTTCAACGCGGCTGTGCTCCATCTGTGAGACTGAC 60
DB 241 ATGTGACAGAGGCGAGAGAGCTTTTCAACGCGGCTGTGCTCCATCTGTGAGACTGAC 300
QY 61 GGCACGCTAAACGCGCAAGCTTACGCGTGTCCGCGAGAGGCGAGGCGATGCCACTTAC 120
DB 301 GGCACGCTAAACGCGCAAGCTTACGCGTGTCCGCGAGAGGCGAGGCGATGCCACTTAC 360
QY 121 GGCAGAGCTGAGACCGGAAAGTTATCTGACACCAACCGGCAAGCTGTCCCTGTGCGCAC 180
DB 361 GGCAGAGCTGAGACCGGAAAGTTATCTGACACCAACCGGCAAGCTGTCCCTGTGCGCAC 420
QY 181 CTGCTGACCAACCTGACCTTACGCGGTGAGAGTTCAGCGCTTACCGGCTACCCGACCATGAG 240
DB 421 CTGCTGACCAACCTGACCTTACGCGGTGAGAGTTCAGCGCTTACCGGCTACCCGACCATGAG 480
QY 241 CAGCAGAGCTTCTTCAAGTCCGCAATGCGCGAGGCTTACGCGAGGCGACCATCTTC 300
DB 481 CAGCAGAGCTTCTTCAAGTCCGCAATGCGCGAGGCTTACGCGAGGCGACCATCTTC 540
QY 301 TTCAAGAGCAGAGGCGACTTACAGACCGCGCGAGAGGTAAGTTTCAGGCGGACACCTG 360
DB 541 TTCAAGAGCAGAGGCGACTTACAGACCGCGCGAGAGGTAAGTTTCAGGCGGACACCTG 600
QY 361 GTGAACCGCATGAGCTGAGAGGCGATGACTTCAAGAGAGCGGCAACATCTGCGGAC 420
DB 601 GTGAACCGCATGAGCTGAGAGGCGATGACTTCAAGAGAGCGGCAACATCTGCGGAC 660
QY 421 AAGCTGAGAGTCACTTACCAAGCCCAAGCGTCTATCTATGAGGCGGCAAGCAAGAAC 480
DB 661 AAGCTGAGAGTCACTTACCAAGCCCAAGCGTCTATCTATGAGGCGGCAAGCAAGAAC 720
QY 481 GGCATCAAGGTGAACTTCAAGATCGCGCAACATCGAGAGCGGACGCTGACCTGCGC 540
DB 721 GGCATCAAGGTGAACTTCAAGATCGCGCAACATCGAGAGCGGACGCTGACCTGCGC 780
QY 541 GACCACTACAGAGAGAACACCCCATCGGCGAGCGGCGGCTGTGCTGCCGACCAACAC 600
DB 781 GACCACTACAGAGAGAACACCCCATCGGCGAGCGGCGGCTGTGCTGCCGACCAACAC 840
QY 601 TACCTGAGCAACCGCTGCGGCTGAGCAAGAACACCCCAAGAGAGCGGCAATGATGTC 660
DB 841 TACCTGAGCAACCGCTGCGGCTGAGCAAGAACACCCCAAGAGAGCGGCAATGATGTC 900
QY 661 CTGCTGAGAGTTCGAGACCGCGCGGATGACTTCTCGGATGAGACGAGCTGTACAAAG 720
DB 901 CTGCTGAGAGTTCGAGACCGCGCGGATGACTTCTCGGATGAGACGAGCTGTACAAAG 960
QY 721 CTTAGCCAGAGCTTCCCGCGCGAGGTGAGAGGAGAGATGATGAGACCGTGGCCATGTCT 780
DB 961 CTTAGCCAGAGCTTCCCGCGCGAGGTGAGAGGAGAGATGATGAGACCGTGGCCATGTCT 1020
QY 781 TGTGCCAGAGAGAGCGGAGTGAACCGTCAACCTGTGAGCTGTGTCTTGTGATGATCAAT 840
DB 1021 TGTGCCAGAGAGAGCGGAGTGAACCGTCAACCTGTGAGCTGTGTCTTGTGATGATCAAT 1080
QY 841 GTGTAG 846
DB 1081 GTGTAG 1086

RESULT 7
AAA92002
ID AAA92002 standard; DNA; 4269 BP.
XX
AC AAA92002;
XX
DT 12-JAN-2001 (first entry)
XX
DE PCRBS-dEGFP reporter construct coding sequence.


```

QY 241 CAGACGACTTCTTCAAGTCCGCGATGCGGAGGCTACGTCCAGAGCGACCATCTTC 300
DB 241 CAGACGACTTCTTCAAGTCCGCGATGCGGAGGCTACGTCCAGAGCGACCATCTTC 300
QY 301 TTCAAGAGCAGCGCACTACAAAGACCCGCGCGAGGTGAAGTTGAGGGCGACACCTTG 360
DB 301 TTCAAGAGCAGCGCACTACAAAGACCCGCGCGAGGTGAAGTTGAGGGCGACACCTTG 360
QY 361 GTGAACCGGATCGAGCTGAAGGGATCGACTTCAAGAGGAGCGGCAATCTGGGGGAC 420
DB 361 GTGAACCGGATCGAGCTGAAGGGATCGACTTCAAGAGGAGCGGCAATCTGGGGGAC 420
QY 421 AAGCTGAGTCACTACAAAGCAGCAGCTATATCATGCGCGAGAGAGAGAC 480
DB 421 AAGCTGAGTCACTACAAAGCAGCAGCTATATCATGCGCGAGAGAGAGAC 480
QY 481 GGCATCAAGGTGAATCTTAAGTCCGCCAACAATCGAGAGCGGAGCTGGCC 540
DB 481 GGCATCAAGGTGAATCTTAAGTCCGCCAACAATCGAGAGCGGAGCTGGCC 540
QY 541 GACCACTACGAGCAAGACCCCATCGGCGAGCGGCGCTGCTGCGCGAGAACAC 600
DB 541 GACCACTACGAGCAAGACCCCATCGGCGAGCGGCGCTGCTGCGCGAGAACAC 600
QY 601 TACCTGAGCAGCCAGTCCGCTGAGCAAGACCCCAACGAGAGCGGATCAATGTC 660
DB 601 TACCTGAGCAGCCAGTCCGCTGAGCAAGACCCCAACGAGAGCGGATCAATGTC 660
QY 661 CTGCTGAGTCTGTGACCGCGCGGAGTCACTCTGCGGATGAGAGACTGTACAAAGAG 720
DB 661 CTGCTGAGTCTGTGACCGCGCGGAGTCACTCTGCGGATGAGAGACTGTACAAAGAG 720
QY 721 CTGCTGAGTCTGTGACCGCGCGGAGTCACTCTGCGGATGAGAGACTGTGCT 780
DB 721 CTGCTGAGTCTGTGACCGCGCGGAGTCACTCTGCGGATGAGAGACTGTGCT 780
QY 781 TGTGCGCAGAGAGCGGAGTGAACCTGCAACCTGAGCCTGTGCTTGTGCTAGATCAAT 840
DB 781 TGTGCGCAGAGAGCGGAGTGAACCTGCAACCTGAGCCTGTGCTTGTGCTAGATCAAT 840
QY 841 GTGTA 845
DB 841 GTGTA 845

RESULT 10
AADS2696 standard; DNA; 10417 BP.
AADS2696;
14-MAY-2003 (first entry)
pmDelcAB-GFP-lcmvNP for in vitro transcription of recombinant replicon.
Non-structural protein; viral replication; vaccine; immunostimulant;
gene therapy; linker; transcription; ds.
Unidentified.
OS
XX Key Location/Qualifiers
XX misc_feature 1
XX FT /tag= a
XX FT /note= "Corresponds to first one of replicon RNA"
XX FT Promoter 10399..10417
XX FT /tag= b
XX FT /note= "T7 promoter"
XX WO200295023-A2.
XX 28-NOV-2002.
XX 23-MAY-2002; 2002WO-IB002810.
PF

```

```

XX 23-MAY-2001; 2001US-029251SP.
XX (INSP) INST PASTEUR.
XX Escriou N, Van Der Werf S, Vignuzzi M, Gerbaud S;
XX WPI; 2003-140373/13.
XX Replicons from a positive strand RNA virus genome for expressing proteins
XX in animal cells, comprises a sequence encoding viral non-structural
XX PT proteins, a sequence for replication, and a sequence encoding a
XX PT heterologous protein.
XX Claim 52; Page 73-76; 76pp; English.
XX
XX The present invention relates to self-replicating recombinant positive
XX strand RNA molecule of a viral genome of an RNA virus, where the molecule
XX comprises an RNA sequence encoding the non-structural proteins of the RNA
XX virus, viral non-encoding RNA sequences necessary for viral replication
XX and an RNA sequence encoding a heterologous protein or a fragment of a
XX heterologous protein. The self-replicating recombinant positive strand
XX RNA molecules are useful in vaccines for expressing heterologous proteins
XX in animal cells and for monitoring RNA replication and RNA delivery into an
XX animal host. They are also in gene therapy. The present sequence is
XX pmDelcAB-GFP-lcmvNP plasmid DNA used for in vitro transcription of
XX recombinant replicons. This sequence is used in the exemplification of
XX the invention
XX
XX Sequence 10417 BP; 2784 A; 2524 C; 2646 G; 2463 T; 0 U; 0 Other;
XX
XX Query Match 84.8%; Score 721.2; DB 7; Length 10417;
XX Best Local Similarity 97.6%; Pred. No. 2.6e-116;
XX Matches 732; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
QY 1 ATGTTAGCAAGGCGGAGAGCTGTCAACCGGGGTGTGCCCATCTGTGAGCTGAGC 60
DB 743 ATGTTAGCAAGGCGGAGAGAGCTGTCAACCGGGGTGTGCCCATCTGTGAGCTGAGC 802
QY 61 GCGCAGTAAACGGGCAACAGTTCAGCGTGTCCGCGAGGCGAGGCGAGTCCACTAC 120
DB 803 GCGCAGTAAACGGGCAACAGTTCAGCGTGTCCGCGAGGCGAGGCGAGTCCACTAC 862
QY 121 GGCAGCTGACCTGAACTTCATCTGCACCAACCGGCAAGCTGCGCTGCGCTCCAC 180
DB 863 GGCAGCTGACCTGAACTTCATCTGCACCAACCGGCAAGCTGCGCTGCGCTCCAC 922
QY 181 CTGTTGACCACTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGAC 240
DB 923 CTGTTGACCACTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGAC 982
QY 241 CAGCAGCACTTCTTCAAGTCCGCGATGCGGAGGCTACGTCCAGAGCGACCATCTTC 300
DB 983 CAGCAGCACTTCTTCAAGTCCGCGATGCGGAGGCTACGTCCAGAGCGACCATCTTC 1042
QY 301 TTCAAGAGCAGCGCACTACAAAGACCCGCGCGAGGTGAAGTTGAGGGCGACACCTTG 360
DB 1043 TTCAAGAGCAGCGCACTACAAAGACCCGCGCGAGGTGAAGTTGAGGGCGACACCTTG 1102
QY 361 GTGAACCGGATCGAGCTGAAGGGATCGACTTCAAGAGGAGCGGCAATCTGGGGGAC 420
DB 1103 GTGAACCGGATCGAGCTGAAGGGATCGACTTCAAGAGGAGCGGCAATCTGGGGGAC 1162
QY 421 AAGCTGAGTCACTACAAAGCAGCAGCTATATCATGCGCGAGAGAGAGAC 480
DB 1163 AAGCTGAGTCACTACAAAGCAGCAGCTATATCATGCGCGAGAGAGAGAC 1222
QY 481 GGCATCAAGGTGAATCTTAAGTCCGCCAACAATCGAGAGCGGAGCTGGCC 540
DB 1223 GGCATCAAGGTGAATCTTAAGTCCGCCAACAATCGAGAGCGGAGCTGGCC 1282
QY 541 GACCACTACGAGCAAGACCCCATCGGCGAGCGGCGCTGCTGCGCGAGAACAC 600

```


DB 1283 GACCACTACGAGACACCCCTATCGCGAGCGCCCGCTGCTGCTCCGACCAACAC 1342
 QY 601 TACCGACACACCCAGTCCGCCCTGACAAAGACCCCAACGAGACGGATCATGTGTC 660
 DB 1343 TACCTGAGACACCCAGTCCGCCCTGACAAAGACCCCAACGAGACGGATCATGTGTC 1402
 QY 661 CTGCTGAGATTGATGACCGCCCGCGGATCACTCTCGCATGAGACGACTGTACAAAG 720
 DB 1403 CTGCTGAGATTGATGACCGCCCGCGGATCACTCTCGCATGAGACGACTGTACAAAG 1462
 QY 721 CTGACCATGCTTCCCGCGGAGTGAG 750
 DB 1463 CTCGAGATGCTCTGCTTAAAGAGATTAA 1492

RESULT 11

ABK86533
 ID ABK86533 standard; DNA; 6418 BP.

AC ABK86533;

XX 07-AUG-2003 (revised)
 DT 30-AUG-2002 (first entry)

XX EIAV based retroviral vector PONY8.1G.

XX Parkinson's disease; gene therapy; retroviral vector; ds; cyclic; EIAV;
 XX neuroprotective; antiparkinsonian; NOI; IRS5; lentiviral particle;
 XX nucleotide site of interest; Internal Ribosome Entry Site; PONY8.1G;
 XX tyrosine hydroxylase; GTP-cyclohydrolase 1; tritestic cassette;
 XX Aromatic Amino Acid Dopa Decarboxylase; tritestic cassette;
 XX Vesicular Monoamine Transporter 2; neurodegenerative disease.

XX Equine infectious anemia virus.
 OS Human cytomegalovirus.
 OS Escherichia coli.
 OS Aequorea victoria.
 OS Synthetic.

XX WO200229065-A2.

XX 11-APR-2002.

XX 05-OCT-2001; 2001WO-GB004433.

XX 06-OCT-2000; 2000GB-00024550.

XX (OXFO-) OXFORD BIOMEDICA UK LTD.

XX Kingsman AU, Mazarakis ND, Martin-Rendon E, Azzouz M, Rohli J;

XX WPI; 2002-507885/54.

XX A novel retroviral vector genome is useful in the treatment of
 PT Parkinson's disease.

XX Example 8; Page 61-62; 106bp; English.

PS The invention relates to a retroviral vector genome (GI) comprising two
 CC or more NOIs (nucleotide sites of interest), operably linked by one or
 CC more Internal Ribosome Entry Site(s) (IRES). Also included are a vector
 CC system (SI) comprising (GI), production (MI) of a lentiviral particle, a
 CC viral particle (PI) produced by (MI) or by (SI), a pharmaceutical
 CC composition comprising (GI), (SI) or (PI), a bistrionic cassette
 CC comprising a nucleotide sequence which encodes: a) tyrosine hydroxylase
 CC and a nucleotide sequence that encodes GTP-cyclohydrolase 1 operably
 CC linked by one or more IRES(s); or b) Aromatic Amino Acid Dopa
 CC Decarboxylase and a nucleotide sequence which encodes Vesicular Monoamine
 CC Transporter 2 operably linked by one or more IRES(s); a tritestic
 CC cassette comprising a nucleotide sequence which encodes tyrosine
 CC hydroxylase, a nucleotide sequence that encodes GTP-cyclohydrolase 1 and
 CC a nucleotide sequence which encodes Aromatic Amino Acid Dopa

CC Decarboxylase, operably linked by one or more IRES(s) and a cell that has
 CC been transfected with (SI). GI, the vector system (SI), and the viral
 CC particle (PI) can all be used to treat and/or prevent (by gene therapy) a
 CC neurodegenerative disease, especially Parkinson's disease in a subject,
 CC and also in the manufacture of a pharmaceutical composition to treat the
 CC above mentioned disease. Treatment with vectors capable of delivering
 CC e.g., Tyrosine hydroxylase, GTP-cyclohydrolase 1, Aromatic Amino Acid
 CC Dopa, Decarboxylase and Vesicular Monoamine Transporter 2, are useful for
 CC the last stages of treatment for sufferers of Parkinson's disease where
 CC they do not respond significantly to L-dopa treatment by prior art
 CC methodologies. The present sequence is retroviral vector of the
 CC invention, PONY8.1G, comprising elements of the EIAV (Equine infectious
 CC anaemia virus) genome, human cytomegalovirus sequences, a plasmid
 CC backbone and a green fluorescent protein sequence. (Updated on 07-AUG-
 CC 2003 to correct OS field.)

XX Sequence 6418 BP; 1673 A; 1559 C; 1570 G; 1616 T; 0 U; 0 Other;

XX Query Match 84.8%; Score 720.6; DB 6; Length 6418;

XX Best Local Similarity 96.8%; Pred. No. 3.3e-116;

XX Matches 735; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 1 ATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTCCATCTGTGAGCTGAC 60
 DB 2690 ATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTCCATCTGTGAGCTGAC 2749
 QY 61 GGGCAAGTAAACGGCCCAAGTTGAGCGTGTCCGCGCAAGGGCGATGCCACTTAC 120
 DB 2750 GGGCAAGTAAACGGCCCAAGTTGAGCGTGTCCGCGCAAGGGCGATGCCACTTAC 2809
 QY 121 GGGCAAGTAAACGGCCCAAGTTGAGCGTGTCCGCGCAAGGGCGATGCCACTTAC 180
 DB 2810 GGGCAAGTAAACGGCCCAAGTTGAGCGTGTCCGCGCAAGGGCGATGCCACTTAC 2869
 QY 181 CTCGTGACCACTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGA 240
 DB 2870 CTCGTGACCACTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGA 2929
 QY 241 CAGACGACCTTCTTAAAGTCCGCAATGCGGAGGCTTGAAGCTTGAAGCTTGAAGCTTGA 300
 DB 2930 CAGACGACCTTCTTAAAGTCCGCAATGCGGAGGCTTGAAGCTTGAAGCTTGAAGCTTGA 2989
 QY 301 TTCAAGGACGAGGAGGAGTAAAGACCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
 DB 2990 TTCAAGGACGAGGAGGAGTAAAGACCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3049
 QY 361 GTGAACCGCATCGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
 DB 3050 GTGAACCGCATCGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3109
 QY 421 AAGCTGAGTACACTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
 DB 3110 AAGCTGAGTACACTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3169
 QY 481 GGCATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540
 DB 3170 GGCATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3229
 QY 541 GACCACTACGACGAGACACCCCATGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
 DB 3230 GACCACTACGACGAGACACCCCATGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3289
 QY 601 TACCTGAGACACCCAGTCCGCCCTGACAAAGACCCCAACGAGACGGATCATGTGTC 660
 DB 3290 TACCTGAGACACCCAGTCCGCCCTGACAAAGACCCCAACGAGACGGATCATGTGTC 3349
 QY 661 CTGCTGAGATTGATGACCGCCCGCGGATCACTCTCGCATGAGACGACTGTACAAAG 720
 DB 3350 CTGCTGAGATTGATGACCGCCCGCGGATCACTCTCGCATGAGACGACTGTACAAAG 3409
 QY 721 CTGACCATGCTTCCCGCGGAGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 759
 DB 3410 ACAGGCGGAGCTTAAAGTGAAGCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3448

CC This invention relates to a novel multicistronic retroviral vector genome
CC comprising a first nucleic acid sequence upstream of an integral
CC regulatory element, such that the level of genomic RNA available for
CC packaging in the absence of rev, or its functional equivalent, is
CC increased. Specifically, the vector is a lentiviral genome that comprises
CC a therapeutically important nucleotide sequence of interest operably
CC linked to a tetracycline operator. Furthermore, the first nucleic acid
CC sequence represents a selectable marker such as the tetracycline
CC repressor (Tet^r) gene, which is codon optimized for expression in
CC mammalian cells. Together, the retroviral vector particle, DNA construct
CC and mammalian cell are useful for the preparation of a medicament to
CC deliver the therapeutic sequence of interest to a target site. As such,
CC the present invention describes a method that can be used in the
CC treatment of inflammatory disease, asthma, psoriasis, cancer e.g. chronic
CC myeloid leukaemia (CML) and neurodegenerative disease e.g. Alzheimer's.
CC Accordingly, via gene therapy these vectors can be described as having
CC virulent, neurotropic, neuroprotective, immunosuppressive and/or
CC antiinflammatory activity. This polynucleotide sequence represents the
CC total plasmid DNA of pONY8.1G of the invention.

XX Sequence 6418 BP; 1673 A; 1559 C; 1570 G; 1616 T; 0 U; 0 Other;

Query Match 84.8%; Score 720.6; DB 8; Length 6418;

Best Local Similarity 96.8%; Pred. No. 3.3e-116;

Matches 735; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 1 ATGCTAGAGCAAGGCGGAGGAGCTGTTCAACGGGGTGGTGGCCATCTCTGGTCAGAGCTGAGAC 60
DB 2690 ATGCTAGAGCAAGGCGGAGGAGCTGTTCAACGGGGTGGTGGCCATCTCTGGTCAGAGCTGAGAC 2749
QY 61 GGGGACGTAAACGGCCCAAGTTCAAGCTGTCGGGAGGCGGAGGCGGAGCTGAGAC 120
DB 2750 GGGGACGTAAACGGCCCAAGTTCAAGCTGTCGGGAGGCGGAGGCGGAGCTGAGAC 2809
QY 121 GGGGACGTAAACGGCCCAAGTTCAAGCTGTCGGGAGGCGGAGGCGGAGCTGAGAC 180
DB 2810 GGGGACGTAAACGGCCCAAGTTCAAGCTGTCGGGAGGCGGAGGCGGAGCTGAGAC 2869
QY 181 CTGCTAGAGCAAGGCGGAGGAGCTGTTCAACGGGGTGGTGGCCATCTCTGGTCAGAGCTGAGAC 240
DB 2870 CTGCTAGAGCAAGGCGGAGGAGCTGTTCAACGGGGTGGTGGCCATCTCTGGTCAGAGCTGAGAC 2929
QY 241 CAGCAGCACTTCTTCAAGTCCGCGCATGCCGAGGCTTCAAGGAGGCGGAGCTGAGAC 300
DB 2930 CAGCAGCACTTCTTCAAGTCCGCGCATGCCGAGGCTTCAAGGAGGCGGAGCTGAGAC 3089
QY 301 TTGAGAGCAAGGCGGAGGAGCTGTTCAACGGGGTGGTGGCCATCTCTGGTCAGAGCTGAGAC 360
DB 2990 TTGAGAGCAAGGCGGAGGAGCTGTTCAACGGGGTGGTGGCCATCTCTGGTCAGAGCTGAGAC 3049
QY 361 GTGAGAGCAAGGCGGAGGAGCTGTTCAACGGGGTGGTGGCCATCTCTGGTCAGAGCTGAGAC 420
DB 3050 GTGAGAGCAAGGCGGAGGAGCTGTTCAACGGGGTGGTGGCCATCTCTGGTCAGAGCTGAGAC 3109
QY 421 AAGCTGAGTAACTTCAAGGAGGAGCTGTTCAACGGGGTGGTGGCCATCTCTGGTCAGAGCTGAGAC 480
DB 3110 AAGCTGAGTAACTTCAAGGAGGAGCTGTTCAACGGGGTGGTGGCCATCTCTGGTCAGAGCTGAGAC 3169
QY 481 GGGATCAAGGAGGAGGAGGAGCTGTTCAACGGGGTGGTGGCCATCTCTGGTCAGAGCTGAGAC 540
DB 3170 GGGATCAAGGAGGAGGAGGAGCTGTTCAACGGGGTGGTGGCCATCTCTGGTCAGAGCTGAGAC 3229
QY 541 GACCACTACGAGCAAGGAGGAGGAGCTGTTCAACGGGGTGGTGGCCATCTCTGGTCAGAGCTGAGAC 600
DB 3230 GACCACTACGAGCAAGGAGGAGGAGCTGTTCAACGGGGTGGTGGCCATCTCTGGTCAGAGCTGAGAC 3289
QY 601 TACCTGAGCAAGGAGGAGGAGGAGCTGTTCAACGGGGTGGTGGCCATCTCTGGTCAGAGCTGAGAC 660
DB 3290 TACCTGAGCAAGGAGGAGGAGGAGCTGTTCAACGGGGTGGTGGCCATCTCTGGTCAGAGCTGAGAC 3349
QY 661 CTGCTGAGTAACTTCAAGGAGGAGGAGCTGTTCAACGGGGTGGTGGCCATCTCTGGTCAGAGCTGAGAC 720
DB 3350 CTGCTGAGTAACTTCAAGGAGGAGGAGCTGTTCAACGGGGTGGTGGCCATCTCTGGTCAGAGCTGAGAC 3409

QY 721 CTGAGCAAGGAGGAGGAGGAGCTGTTCAACGGGGTGGTGGCCATCTCTGGTCAGAGCTGAGAC 759
DB 3410 AGCGGCGGAGGAGGAGGAGGAGCTGTTCAACGGGGTGGTGGCCATCTCTGGTCAGAGCTGAGAC 3448

RESULT 14

ID ABS57536/c

ABS57536 standard; cDNA; 12789 BP.

AC ABS57536;

DT 10-FEB-2003 (first entry)

DE Vector pK7MG2D DNA.

KW PDR1; pleiotropic drug resistance; ATP-binding cassette-transporter;

KW ABC-transporter; secondary metabolite production; transgenic plant;

KW plant disease resistance; secondary metabolite secretion; plant;

KW pest resistance; anti-oxidative; anti-metastatic; alkaloid; cytostatic;

KW terpenoid indole alkaloid tropone alkaloid; steroid alkaloid;

XX polyhydroxy alkaloid; pesticide; vector; ds.

XX Synthetic.

PN WO20028388-A2.

PP 24-OCT-2002.

PR 18-APR-2002; 2002MO-EP004322.

PR 18-APR-2001; 2001EP-00201407.

PA (VLA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.

PI Goossens A, Inze D;

DR WPI, 2003-058637/05.

XX Use of expression cassette comprising gene encoding ATP-binding cassette-

XX transporter to induce or enhance production or secretion of secondary

XX metabolite by plant or plant cells.

XX Example 3; Page 49-55; 58pp; English.

XX This invention describes the novel use of an expression cassette

XX comprising a gene encoding an ATP-binding cassette (ABC)-transporter to

XX induce/enhance production/secretion of secondary metabolite by plant

XX cells, or to stimulate production of the secondary metabolite by plants.

XX The method comprises transforming plant cells or the plant with the

XX expression vector containing the expression cassette, selecting the

XX transformed transgenic plant or plant cells, and propagating them. The

XX expression cassette is useful for inducing or enhancing production or

XX secretion of secondary metabolites in plants or plant cells, which e.g.

XX play a role in the resistance of plants to pests and diseases and which

XX can have anti-oxidative and anti-metastatic properties. The secondary

XX metabolites are preferably alkaloids e.g. terpenoid indole alkaloids. The

XX tropone alkaloids, steroid alkaloids and polyhydroxy alkaloids. The

XX cassette of the invention has cytostatic and pesticidal activity. This

XX sequence represents the vector pK7MG2D which is used in the expression of

XX the pleiotropic drug resistance genes described in the disclosure of the

XX invention

XX Sequence 12789 BP; 3050 A; 3326 C; 3397 G; 3015 T; 0 U; 1 Other;

XX Query Match 84.7%; Score 720; DB 7; Length 12789;

XX Best Local Similarity 100.0%; Pred. No. 4.2e-116;

XX Matches 720; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTAGAGCAAGGCGGAGGAGGAGCTGTTCAACGGGGTGGTGGCCATCTCTGGTCAGAGCTGAGAC 60

DB 5797 ATGCTAGAGCAAGGCGGAGGAGGAGGAGCTGTTCAACGGGGTGGTGGCCATCTCTGGTCAGAGCTGAGAC 5738

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2004, 02:53:31 ; Search time 2816 Seconds

(without alignments)
9013.797 Million cell updates/sec

Title: US-09-931-232-2

Perfect score: 850

Sequence: 1 atggtgagcgaagcgagga.....tagatcaatgtagatgc 850

Scoring table: IDENTITY NUC

Gapop 10.0 ; Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
EST:
1: em_estba:*
2: em_esthm:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hcc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hcc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrl:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	515.2	60.6	697	13	BU626888 UI-H-FTO-
2	493.8	58.1	790	13	AL044652 DKFZP434P
3	456.4	53.7	635	14	CD369111 UI-H-FTI-
4	438.4	51.6	615	12	BM209799 C0653E03-

Result No.	Score	Query Match	Length	ID	Description
5	417.4	49.1	536	14	CB529588 UI-H-FT2-
6	411.4	48.4	550	13	BQ551001 H4006G01-
7	274.8	32.3	456	12	BM210789 C0670D12-
8	273	32.1	467	12	BQ034019 UI-1-BY0-
9	234.4	27.6	433	14	CA431885 UI-H-FTO-
10	210.4	24.8	397	9	AJ437938
11	205.2	24.1	317	9	AJ437931
12	205	24.1	511	9	AJ437840
13	168	19.8	570	12	BM194666 EST RERE
14	161.8	19.0	240	12	BM688204 TMM102 Hu
15	161	18.9	525	12	BM736033 TMM102 Hu
16	146	17.2	193	9	AJ437933
17	142	16.7	317	13	BQ551000 H4006G01-
18	130.6	15.4	506	10	BF784326 602110101
19	129.8	15.3	428	12	B1965250 i435G04.Y
20	129.8	15.3	465	12	BM205695 C0287D12-
21	129.8	15.3	467	10	BE629419 u437903.Y
22	129.8	15.3	473	10	BE631772 u46411.Y
23	129.8	15.3	473	10	BE861518 UI-M-AH1-
24	129.8	15.3	481	12	BM246643 K0741A07-
25	129.8	15.3	487	14	CF981920 ma175d02.
26	129.8	15.3	498	14	CB524049 UI-M-GK0-
27	129.8	15.3	504	12	BM249605 K0836R06-
28	129.8	15.3	518	13	BU605375 mdh90F10.
29	129.8	15.3	519	9	A1848698 UI-M-AH1-
30	129.8	15.3	526	10	AW541064 C0140F03-
31	129.8	15.3	548	14	CF949990 UI-M-HK0-
32	129.8	15.3	550	13	BQ930088 AGENCCURT
33	129.8	15.3	555	12	BM054167 i451d10.Y
34	129.8	15.3	562	12	BG863417 602797075
35	129.8	15.3	562	12	BM241829 K0623B12-
36	129.8	15.3	564	13	BQ565843 g146112.Y
37	129.8	15.3	566	9	A1429802 mg04b08.x
38	129.8	15.3	568	12	BM241679 K0620E06-
39	129.8	15.3	570	10	BG055221 H1028R05-
40	129.8	15.3	572	10	AW540851 C0137H08-
41	129.8	15.3	586	12	BG082647 H3078F02-
42	129.8	15.3	590	10	BE630086 u415B05.x
43	129.8	15.3	591	12	BM231966 K0316G05-
44	129.8	15.3	594	14	W98488 mg20911.r1
45	129.8	15.3	609	12	BM241264 K0614E12-

ALIGNMENTS

RESULT 1
BU626888/c 697 bp mRNA linear EST 23-SEP-2002
LOCUS
DEFINITION
UI-H-FTO-bhn-d-01-0-UI.s1 NCI CGAP FTO Homo sapiens CDNA clone
UI-H-FTO-bhn-d-01-0-UI 3', mRNA sequence.
ACCESSION
BU626888
VERSION
BU626888.1 GI:2293103
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 697)
NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
AUTHORS
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Unpublished (1997)
TITLE
Tumor Gene Index
JOURNAL
Contact: Robert Strausberg, Ph.D.
COMMENT
Email: cgaps-remail.nih.gov
Tissue Procurement: Robeff-Panella, U of I
cDNA library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@iowa.edu
Seq primer: ML3 FORWARD
POLYA=Yes.

FEATURES

source

Location/Qualifiers

1..697

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="UT-H-FT0-Dm-4-01-0-UT"

/issue_type="Alveolar Macrophage"

/dev_stage="Adult"

/lab_host="DH10B (Life Technologies)"

/clone_lib="NCI_CGAP_FTO"

/note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP_FTO is a cDNA library constructed from a pool of 81 RNA samples from Alveolar Macrophages challenged with different treatments. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is GGCCATGCCG. The cell line was provided by Gary W. Hunninghake from the University of Iowa."

TAG TISSUE=Human Lung Alveolar Macrophage

TAG LIB=UI-H-FT0

TAG_SEQ=GGCCATGCCG"

Query Match

Best Local Similarity 98.6%; Score 515.2; DB 13; Length 697;

Matches 520; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

200 AGGGGCTGAGGCTTACGCGGCTACCCGACCATATGAGAGAGAGAGCTTCTTCAAT 259

697 AGGGGCTGAGGCTTACGCGGCTACCCGACCATATGAGAGAGAGAGCTTCTTCAAT 638

260 CGGCATGCGCCAGGCTACGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 319

637 CGGCATGCGCCAGGCTACGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 578

320 ACAAG 379

577 ACAAG 518

380 AGGGGCTGAGGCTTACGCGGCTACCCGACCATATGAGAGAGAGAGAGAGAGAGAGAG 439

517 AGGGGCTGAGGCTTACGCGGCTACCCGACCATATGAGAGAGAGAGAGAGAGAGAGAG 458

440 ACAGGAG 499

457 ACAGGAG 398

500 AGATCCGACCAACATGAG 559

397 AGATCCGACCAACATGAG 338

560 CCCCCATGCGGAG 619

337 CCCCCATGCGGAG 278

620 CCGTGAAG 679

277 CCGTGAAG 218

680 CCGGCGGAGATCACTTCGGGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 727

217 CCGGCGGAGATCACTTCGGGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 170

RESULT 2

AL044652

LOCUS 790 bp mRNA linear EST 04-SEP-2003

DEFINITION DKFZp434P092.r1.434 (synonym: htes3) Homo sapiens cDNA clone DKFZp434P092.5, mRNA sequence.

ACCESSION AL044652

VERSION AL044652.1 GI:5432867

KEYWORDS EST.

SOURCE

ORGANISM Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

Ausio, J., Bensa, V., Krieger, S., Mewes, H.W., Gassenhuber, J. and Wiemann, S.

EST (Ausio, Bensa, et al.)

Unpublished (1999)

CONTACT: MIPS

MIPS

Ingo Isenhardt, Landstr. 1, D-85764 Neuberg, Germany

This is the 5' sequence of the clone insert

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ). Email: s.wiemann@dkfz-heidelberg.de; Heidelberg/Germany) within the cDNA sequencing consortium of the German Genome Project.

No sl sequence available.

This clone (DKFZp434P092) is available at the RZPD in Berlin. Please contact the RZPD: Resourcenentrum, Heubergweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

source

Location/Qualifiers

1..790

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="DKFZp434P092"

/issue_type="testis"

/dev_stage="adult"

/lab_host="DH10B"

/note="Vector: pSPORT1; Site 1: NotI; Site 2: SalI"

Query Match

Best Local Similarity 58.1%; Score 493.8; DB 9; Length 790;

Matches 578; Conservative 0; Mismatches 122; Indels 1; Gaps 1;

17 AGGAGCTGTGACCGGGGTGATCCATCTGCTGAGCTGAGAGAGAGAGAGAGAGAGAG 76

1 AGGAGCTGTGACCGGGGTGATCCATCTGCTGAGCTGAGAGAGAGAGAGAGAGAGAG 59

77 ACAAGTTCAGGCTGTCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 136

60 ACAAGTTCAGGCTGTCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 119

137 AGTTCATCTGACCAACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 196

120 AATTCATCTGACCAACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 179

197 CCTAGGAGTGAAGGCTTACGCGGCTACCCGACCATATGAGAGAGAGAGAGAGAGAGAG 256

180 CCTAGGAGTGAAGGCTTACGCGGCTACCCGACCATATGAGAGAGAGAGAGAGAGAGAG 239

257 AGTCCGATGCGCGAG 316

240 AGAGGCGCATGCGCGAG 299

317 ACTAAG 376

300 ACTAAG 359

377 TGAAGGAGATGAGCTTCAAG 436

360 TGAAGGAGATGAGCTTCAAG 419

QY	437	ACAACAGCGCAACAAGTCTATATCAATGGCCGACAGACAGAAAGACCGCATCAAGGTGA	496
Db	420	ATAACTCCCAACAATGTGTACATTCATGGCGCAACAACAAAAGATGGCATCAAGGTCA	479
QY	497	TCAGATTCGGCCACAACATTCGAGAGACGGCAGCGTGCAGTCCGCGACCACTACCGACGA	556
Db	480	TCAGATCAGACACAACAATGTAGAGATGATCGTGCAGCTGGCCACACATTATCAACGA	539
QY	557	ACACCCCATCGGCGACGAGCCCGTGTGCTCCGACACACACTACCTGACGACCCAGT	616
Db	540	AACATCCGATGGGCGACGGCCCTGTGTCTCTCCGACACACATTACTGTCCACCCAGT	599
QY	617	CCGCCCTGAGCAAGACCCCAAAGAAAGCGGATCACTGCTCTGTGAGTTCTGA	676
Db	600	CTGCCCTGTCTAAATATCCCAACGAAAAGAGACCAATGCTCTGTGAGTTGTGA	659
QY	677	CCGCGCGCGGAGTCACTCTCGGCAATGAGCGAGCTTACAG	717
Db	660	CCGCTGTGGGATCACACATGGCAATGAGCGAGCTTTCAG	700

RESULT 3	CD369111/c	LOCUS	DEFINITION	CD369111	635 bp	mRNA	linear	EST 29-MAY-2000
				UH-H-FT1-bj-y-m-07-0-UI	s1	NCI CGAP_F11 Homo sapiens	CDNA clone	
				UH-H-FT1-bj-y-m-07-0-UI	3'	mRNA sequence.		

ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
CB369111	CB369111.1	EST.	Homo sapiens	Homo sapiens (human)

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. Gary W. Hunninghake, U of I
CNA Library preparation: Dr. M. Bento Soares, University of Iowa
DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/cgab.html>
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES	Location/Qualifiers
source	1. .635

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-FT1-bjy-m-07-0-UI"
/tissue_type="Alveolar Macrophage"
/dev_stage="Adult"
/lab_host="RDH10B (Life Technologies)"
/clone_idb="NCI_CGAP_FTI"
/notes="Organ: Lung; Vector: pPT73-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
NCI_CGAP_FTI is a normalized cDNA library constructed from
a pool of 81 RNA samples from Alveolar Macrophages
challenged with different treatments. The library was
normalized according to Bonaldo, Lennon and Soares, Genomics
Research, 6:791-806, 1996. First strand cDNA synthesis was
primed with an oligo-dT primer containing a Not I site.
Double stranded cDNA was ligated to an EcoR I adaptor,
digested with Not I, and cloned directionally into
pPT73-Pac vector. The oligonucleotide used to prime the
synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is

```

GGCATGCGC. The tissue was provided by Dr. Gary W.
 Hunninghake of the University of Iowa.
 TAG_TISSUE=Human Lung Alveolar Macrophage
 TAG_LIB=01-H-FT1
 TAG_SEQ=GGCATGCGC"

```

Query Match      53.7%  Score 456.4  DB 14  Length 635;
Best Local Similarity 99.8%  Pred. No. 2.5-52;
Matches 457, Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

262 GCCATGCCCGAAGCTACGTCCAGGAGGCCATCTTCTTCAAGGACGACGCACTAC 321

Db 635 GCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGACGACGGCAACTAC 576

Qy 322 AAGACCGCGCCGAGGTGAATTGAGGGGACACCTTGTTAAACCGCATCGAGCTGAAG 381

Db 575 AAGACCCCGCGCCGAGGTGAATTGAGGGCGACACCTTGTTGAACCGCATCCGACTGAAG 516

382 GGCATCGACTTCAAGGAGGACGGCAACATCTCTGGGGGACACAAAGCTGGAGTACAACTAAC 441

[illegible]

445 AGGCCCAACCGTATATCATGGCGACAGACAGAGAACGGCATCAAGGTGAATTCTTCAAG 395
47 AGGCCCAACCGTATATCATGGCGACAGACAGAGAACGGCATCAAGGTGAATTCTTCAAG 392

502 ATTCGCCCAACAATTCAGGACGGCAGCGTGCAGCTGCCGCCACCACTACGACAGAACACC 561

Db 395 ATCCGCCAACAATCGAGAGCGCAGCGTGCAGCTCCCGACCACTACGAGCAAGACAC 336

562 CCCATCGGCAGCGCCCGTGCTGCTGCGCGACAAACAATCTGAGCACCAGTCCGCC 621

Dd 335 CCCATCGGCGACGGCCCCCGTGTCTGTGCTGCGACCACTACTGAGGACCCAGTCCGCC 276

622 CTGAGCAAGACCCCAACGAGAGCGGCATCATGTCTCTGCTGAGTTCTGTACCGCC 681

D**b** 275 CTGAGCAAGACCCCAAGAGCGCATCAATGTCTCTGGAGTTCGTACCGCC 216

682 GCCGGGATCACTCTGGCATGGACGAGCTGTACAGAA 719

Db 215 GCCGGATCACTCTCGSCATGACGAGCTGTACAAGTA 178

RESULT 4
DM200700/2

LOCUS	615 bp	mRNA	linear	EST	08-JUN-2003
DEFINITION					
C0653E03-3	NIA Mouse Trophoblast Stem Cell	cdna library (Short)	MUS		

ACCESSION	musculus cDNA clone C0653E03 3', mRNA sequence.
BM209799	

```

VERSION      BM209799.2  GI:31533410
KEYWORDS     EST.

```

SOURCE ORGANISM	Mus musculus (house mouse)	Mus musculus
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE	AUTHORS	TITLE
1 (Tables 1 to 615)	Piao, Y., Ko, N. T., Lim, M. K. and Ko, M. S. H.	Construction of long-transcript enriched cDNA libraries from

method
subprogram
amounts of total RNAs by a universal PCR amplification
method

JOURNAL
Genome Res. 11 (9), 1553-1558 (2001)
MEDLINE
21429098

PUBMED
11544139
On Dec 14, 2001 this sequence version replaced gi:17766404.
COMMENT

Contact: Dawood B. Dudaekula
 Laboratory of Genetics
 National Institute on Aging/National Institutes of Health

Medical Institute on Aging/National Institute on
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@iasun.arc.nia.nih.gov

```

Plate: C0653 row: E column: 03
Seq primer: -21M13 Forward

```

High quality sequence stop: 615

FEATURES
source

POLYA=Yes.

Location/Qualifiers

1. 615
/organism="Mus musculus"
/mol_type="mRNA"
/strain="B6/EGFP transgenic ICR mice"
/db_xref="taxon:10090"
/clone="C0653E03"
/issue_type="Trophoblast stem cell"
/dev_stage="3.5-dpc"
/lab_host="DH10B"
/clone_lib="NIA Mouse Trophoblast Stem Cell cDNA Library (Short)"
/note="Vector: pSPORT1 (Invitrogen); Site 1: SalI; Site 2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA). This is a short-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were obtained from Dr. Janet Rossant and Tilo Kunath (Samuel Lunenfeld Research Institute, Canada). Double-stranded cDNAs were synthesized with an Oligo(dT) primer
[Invitrogen: 5'-pGACTAGTCTAGATCGGAGCGCCGCTTTT-3'] from 4 ug of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to lome-linker L1-SalI, purified by phenol/chloroform, and separated from free linkers by centrifugation 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer SalI-L. The products were purified by phenol/chloroform and centrifuged 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 1.3 Kb. The library was constructed by Yulan Piao (NIA)."

ORIGIN

Query Match 51.6%; Score 438.4; DB 12; Length 615;
Best Local Similarity 86.4%; Pred. No. 6.6e-50;
Matches 484; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

279 CGTCCAGAGGCGGACCATCTTTCTTCAAGAGCGGCACTACAGACCGCGCCGAGGT 338
615 CGTCCAGAGGCGGACCATCTTTCTTCAAGAGCGGCACTACAGACCGCGCCGAGGT 556
339 GAAGTTCGAGGGGAGCAACCTGTGTAACCGCATCGAGCTGAGGGCATGCACTTCAAGGA 398
555 GAAGTTCGAGGGGAGCAACCTGTGTAACCGCATCGAGCTGAGGGCATGCACTTCAAGGA 496
399 GAGCGGCAACATCTCTGGGGGCAAGAGTGAAGTCACTCAACAGGCAACAGCTATAT 458
495 GAGCGGCAACATCTCTGGGGGCAAGAGTGAAGTCACTCAACAGGCAACAGCTATAT 436
459 CATGGCCGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 518
435 CATGGCCGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 376
519 GAGCGGCAACATCTCTGGGGGCAAGAGTGAAGTCACTCAACAGGCAACAGCTATAT 578
375 GAGCGGCAACATCTCTGGGGGCAAGAGTGAAGTCACTCAACAGGCAACAGCTATAT 316
579 CGTGGCTGGCCGAGCAACATCTCTGGGGGCAAGAGTGAAGTCACTCAACAGGCAACAGCT 638
315 CGTGGCTGGCCGAGCAACATCTCTGGGGGCAAGAGTGAAGTCACTCAACAGGCAACAGCT 256
639 CGAGAGCGGAGATCATGTGCTGCTGAGATTGTAACCGCGCGGGATCATCTTCCG 698
255 CGAGAGCGGAGATCATGTGCTGCTGAGATTGTAACCGCGCGGGATCATCTTCCG 196
699 CATGAGCAGAGCTGTACAGAGAGCTTAGCCATGGCTTCCGCGGAGGTGAGAGAGAGGA 758

DB 195 CATGAGCAGAGCTGTACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 136
QY 759 TGATGACAGCGTGGCCAGATGCTGTGGCCAGGAGAGCGGAGATGAGACCGTCAACCCGAGC 818
DB 135 GGTGGCTGGTGTGGCCAGATGCTGTGGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 76
QY 819 CTGTGCTGCTGCTGAGATCA 838
DB 75 CAAAATTAATGAGGAGATCA 56

RESULT 5
CB529588/c 596 bp mRNA linear EST 16-MAY-2003
LOCUS
DEFINITION
UI-H-FRT2-bjh-d-17-0-UI-s1 NCI CGAP_FRT2 Homo sapiens cDNA clone
UI-H-FRT2-bjh-d-17-0-UI 3', mRNA sequence.
CB529588
CB529588.1 GI:29390425
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 596)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapdb-remail.nih.gov
Tissue Procurement: Dr. Gary W. Hunninghake, U of I
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/cgap.html
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES
source

Location/Qualifiers

1. 596
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-FRT2-bjh-d-17-0-UI"
/issue_type="Alveolar Macrophage"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI-CGAP_FRT2"
/note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP_FRT2 is a subtracted cDNA library constructed from a pool of 81 RNA samples from Alveolar Macrophages challenged with different treatments. The library was subtracted according to Bonaldi, Lemon and Soares, Genome Research, 6:791-806, 1996. The tissue was provided by Dr. Gary W. Hunninghake of the University of Iowa.
TAG_TISSUE=Human Lung Alveolar Macrophage
TAG_LIB=UI-H-FRT2
TAG_SEQ=GCCATGCCG"

ORIGIN

Query Match 49.1%; Score 417.4; DB 14; Length 596;
Best Local Similarity 99.8%; Pred. No. 4.5e-47;
Matches 418; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

301 TTCAAGAGAGAGGAGCACTACAGAGCCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
DB 596 TTCAAGAGAGAGGAGCACTACAGAGCCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 537
QY 361 GTGAACCGCATGAGCTGAGAGGGCATGCACTTCAAGAGAGAGGAGGAGGAGGAGGAGGAG 420
DB 536 GTGAACCGCATGAGCTGAGAGGGCATGCACTTCAAGAGAGAGGAGGAGGAGGAGGAGGAG 477

Qy 421 AAGCTGAGTACAACTACAAAGCAAGCAACGTCATATCATGCGCCAGCAAGCAAGAAC 480
 Db 476 AAGCTGAGTACAACTACAAAGCAAGCAACGTCATATCATGCGCCAGCAAGCAAGAAC 417
 Qy 481 GGCATCAAGGTGAATTTCAAGATCCGCCCAACATCGAGAGCGGAGCGTGCAGCTCGCC 540
 Db 416 GGCATCAAGGTGAATTTCAAGATCCGCCCAACATCGAGAGCGGAGCGTGCAGCTCGCC 357
 Qy 541 GACCACTACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 600
 Db 356 GACCACTACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 297
 Qy 601 TACCTGACCACTACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 660
 Db 296 TACCTGACCACTACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 237
 Qy 661 CTGCTGAGTTCGTGACCG 719
 Db 236 CTGCTGAGTTCGTGACCG 178

RESULT 6
 BO551001 550 bp mRNA linear EST 20-JUN-2002
 LOCUS H4006G01-5 NIA Mouse 7.4K cDNA Clone Set Mus musculus cDNA clone
 H4006G01 5', mRNA sequence.
 BO551001
 ACCESSION BO551001
 VERSION BO551001.1 GI:21451887
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE
 AUTHORS VanBuren V., Piao Y., Dudekula D.B., Qian Y., Carter M.G.,
 Martin P.R., Stagg C.A., Bassey U., Alba K., Hamatani T.,
 Kargul G.J., Luo A.G., Kelso J., Hide W. and Ko M.S.H.
 Assembly, verification, and initial annotation of NIA 7.4K mouse
 cDNA clone set
 JOURNAL Genome Res. 12 (12), 1999-2003 (2002)
 MEDLINE 22354164
 PUBMED 12466305

COMMENT
 Other ESTs: H400G01-3
 Contact: Yong Qian
 Laboratory of Genetics
 National Institute on Aging/National Institutes of Health
 333 Cassell Drive, Suite 3000, Baltimore, MD 21224-6820, USA
 Email: cdna@igsun.grc.nia.nih.gov
 This clone set has been freely distributed to the community. Please
 visit http://igsun.grc.nia.nih.gov/cDNA/NIA_7_4K.html for details.
 Plate: H4006 row: G column: 01
 Seq primer: -21M13 Reverse
 High quality sequence stop: 550
 POLYA=No.

FEATURES
 source
 1..550
 /location/Qualifiers
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /db_xref="taxon:10090"
 /clone="H4006G01"
 /sex="mixed"
 /dev_stage="mixed"
 /lab_host="DH10B"
 /clone_lib="NIA Mouse 7.4K cDNA Clone Set"
 /note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; This
 clone is among a rearranged set of 7,407 clones from more
 than 20 cDNA libraries."

ORIGIN
 Query Match 48.4%; Score 411.4; DB 13; Length 550;

Best Local Similarity 99.8%; Pred. No. 2.9e-46;
 Matches 412; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 ATGTGAGCAAGGCGGAGAGCTTTCAACGGGCTGTGCTCCATCTGTGAGCTGAC 60
 Db 138 ATGTGAGCAAGGCGGAGAGCTTTCAACGGGCTGTGCTCCATCTGTGAGCTGAC 197
 Qy 61 GGCAGCTAAACGGCGCAAGTTGAGGCTGTGCGGCGAGGCGGAGGCGAGTGCACCTAC 120
 Db 198 GGCAGCTAAACGGCGCAAGTTGAGGCTGTGCGGCGAGGCGGAGGCGAGTGCACCTAC 257
 Qy 121 GGCAGCTAAACGGCGCAAGTTGAGGCTGTGCGGCGAGGCGGAGTGCACCTGCGCAC 180
 Db 258 GGCAGCTAAACGGCGCAAGTTGAGGCTGTGCGGCGAGGCGGAGTGCACCTGCGCAC 317
 Qy 181 CTGTGACCACTGACCTTACGCGGTGAGTGTCTTACGCGGCTTACCCGACCAATGAG 240
 Db 318 CTGTGACCACTGACCTTACGCGGTGAGTGTCTTACGCGGCTTACCCGACCAATGAG 377
 Qy 241 CAGCAGCACTTCTTCAAGTCGCGGCGGCGGAGGCTTACGCGGAGGCGGAGGCGGAGG 300
 Db 378 CAGCAGCACTTCTTCAAGTCGCGGCGGCGGAGGCTTACGCGGAGGCGGAGGCGGAGG 437
 Qy 301 TTCAAGAGCAAGGCGCACTAAGAGCCGCGCGCGAGGTGAAGTTTCAAGGCGGAGGCGG 360
 Db 438 TTCAAGAGCAAGGCGCACTAAGAGCCGCGCGCGAGGTGAAGTTTCAAGGCGGAGGCGG 497
 Qy 361 GTGAACCGCATCGAGCTGAAGGCGATGACTTCAAGAGAGGCGGACATCTT 413
 Db 498 GTGAACCGCATCGAGCTGAAGGCGATGACTTCAAGAGAGGCGGACATCTT 550

RESULT 7
 LOCUS BM210789/c 456 bp mRNA linear EST 08-JUN-2003
 DEFINITION C0670D12-3 NIA Mouse Trophoblast Stem Cell cDNA Library (Short) Mus
 musculus cDNA clone C0670D12 3', mRNA sequence.
 ACCESSION BM210789
 VERSION BM210789.2 GI:31534239
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE
 AUTHORS Bukayota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 456)
 TITLE Piao Y., Ko N.T., Lim M.K. and Ko M.S.H.
 Construction of long-transcript enriched cDNA libraries from
 submicrogram amounts of total RNAs by a universal PCR amplification
 method
 JOURNAL Genome Res. 11 (9), 1553-1558 (2001)
 MEDLINE 21429098
 PUBMED 11544199

COMMENT
 On Dec 14, 2001 this sequence version replaced gi:17767758.
 Contact: Dawood B. Dudekula
 Laboratory of Genetics
 National Institute on Aging/National Institutes of Health
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
 Email: cdna@igsun.grc.nia.nih.gov
 Plate: C0670 row: D column: 12
 Seq primer: -21M13 Forward
 High quality sequence stop: 456
 POLYA=Yes.

FEATURES
 source
 1..456
 /location/Qualifiers
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="B6/EGFP transgenic ICR mice"
 /db_xref="taxon:C0670D12-3"
 /db_xref="taxon:10090"
 /clone="C0670D12"
 /tissue_type="Trophoblast stem cell"
 /dev_stage="3.5-dpc"
 /lab_host="DH10B"

RESULT 11
AJ437931 317 bp mRNA linear EST 15-MAR-2002
LOCUS AJ437931
DEFINITION Helianthus annuus cDNA clone HADpLR259A02, mRNA
ACCESSION AJ437931
VERSION AJ437931.1 GI:19525470
KEYWORDS EST.
SOURCE Helianthus annuus (common sunflower)
ORGANISM Helianthus annuus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; campanulids; Asterales; Asteraceae; Asteroideae;
Heliantheae; Helianthus.
REFERENCE 1 (bases 1 to 317)
Tamborindegy, C.
An expressed-sequenced-tags database of the sunflower protoplast
Thesis (2002) Department of Ecole Nationale Supérieure Agronomique
de Toulouse, Institut National Polytechnique de Toulouse, Toulouse,
France
JOURNAL Unpublished (2002)
COMMENT Contact: Gentzbitel L
Laboratoire de Biotechnologie et Amélioration des Plantes
Institut National Polytechnique de Toulouse - Ecole National
Supérieure Agronomique de Toulouse
IFR40, Pole de Biotechnologie Vegetale, 18 chemin de Borde Rouge,
Auzeville, CASTANET TOLOSAN 31326, France.
ORIGIN Location/Qualifiers
source
1. 317
/organism="Helianthus annuus"
/mol_type="mRNA"
/cultivar="Emil"
/db_xref="taxon:4232"
/clone="HADpLR259A02"
/issue_type="hypocotyls"
/cell_type="protoplasts"
/dev_stage="1-5 days"
/clone_11b="HADpLR2"
Query Match 24.1%; Score 205.2; DB 9; Length 317;
Best Local Similarity 98.6%; Pred. No. 1.9e-18;
Matches 207; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 ATGTGAGCAAGGCGAGAGCTGTCACCGGGGNGGCCATCTGCTGAGCTGGAC 60
DB 68 ATGTGAGCAAGGCGAGAGCTGTCACCGGGGNGGCCATCTGCTGAGCTGGAC 127
QY 61 GCGCAGTAAACGGCCCAAGTTCAAGCTGTCCGGCGAGGCGAGGCGAGTCCACCTAC 120
DB 128 GCGCAGTAAACGGCCCAAGTTCAAGCTGTCCGGCGAGGCGAGGCGAGTCCACCTAC 187
QY 121 GCGCAGTAAACGGCCCAAGTTCAAGCTGTCCGGCGAGGCGAGGCGAGTCCACCTAC 180
DB 188 GCGCAGTAAACGGCCCAAGTTCAAGCTGTCCGGCGAGGCGAGGCGAGTCCACCTAC 247
QY 181 CTCTGACCACTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGAC 210
DB 248 CTCTGACCACTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGAC 277
RESULT 12
AJ437840 511 bp mRNA linear EST 15-MAR-2002
LOCUS AJ437840
DEFINITION Helianthus annuus cDNA clone HADpLR205H12, mRNA
ACCESSION AJ437840
VERSION AJ437840.1 GI:19525379
KEYWORDS EST.
SOURCE Helianthus annuus (common sunflower)
ORGANISM Helianthus annuus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

REFERENCE
AUTHORS 1 (bases 1 to 511)
Tamborindegy, C.
An expressed-sequenced-tags database of the sunflower protoplast
Thesis (2002) Department of Ecole Nationale Supérieure Agronomique
de Toulouse, Institut National Polytechnique de Toulouse, Toulouse,
France
JOURNAL Unpublished (2002)
COMMENT Contact: Gentzbitel L
Laboratoire de Biotechnologie et Amélioration des Plantes
Institut National Polytechnique de Toulouse - Ecole National
Supérieure Agronomique de Toulouse
IFR40, Pole de Biotechnologie Vegetale, 18 chemin de Borde Rouge,
Auzeville, CASTANET TOLOSAN 31326, France.
ORIGIN Location/Qualifiers
source
1. 511
/organism="Helianthus annuus"
/mol_type="mRNA"
/cultivar="Emil"
/db_xref="taxon:4232"
/clone="HADpLR205H12"
/issue_type="hypocotyls"
/cell_type="protoplasts"
/dev_stage="1-5 days"
/clone_11b="HADpLR2"
Query Match 24.1%; Score 205; DB 9; Length 511;
Best Local Similarity 100.0%; Pred. No. 2.2e-18;
Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTGAGCAAGGCGAGAGCTGTCACCGGGGNGGCCATCTGCTGAGCTGGAC 60
DB 262 ATGTGAGCAAGGCGAGAGCTGTCACCGGGGNGGCCATCTGCTGAGCTGGAC 321
QY 61 GCGCAGTAAACGGCCCAAGTTCAAGCTGTCCGGCGAGGCGAGGCGAGTCCACCTAC 120
DB 322 GCGCAGTAAACGGCCCAAGTTCAAGCTGTCCGGCGAGGCGAGGCGAGTCCACCTAC 381
QY 121 GCGCAGTAAACGGCCCAAGTTCAAGCTGTCCGGCGAGGCGAGGCGAGTCCACCTAC 180
DB 382 GCGCAGTAAACGGCCCAAGTTCAAGCTGTCCGGCGAGGCGAGGCGAGTCCACCTAC 441
QY 181 CTCTGACCACTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGAC 205
DB 442 CTCTGACCACTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGAC 466
RESULT 13
BM194666 570 bp mRNA linear EST 13-DEC-2002
LOCUS BM194666
DEFINITION EST RBRB sh E3+ SSH Library from Adult Cirrhotic Liver Homo sapiens
cDNA, mRNA sequence.
ACCESSION BM194666
VERSION BM194666.1 GI:2665847
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 570)
Shackel, N.A., McGunness, P., Abbott, A., Gorrell, G. and McCaughan, G.
Suppression subtractive hybridisation analysis of intrahepatic
differential gene expression in human liver cirrhosis
Unpublished (2002)
JOURNAL Unpublished (2002)
COMMENT Contact: Shackel NA
Liver Immunobiology Laboratory
Centenary Institute
Locked Bag No. 6, Newtown, NSW 2042, Sydney, Australia
Tel: 61 2 95656101
Fax: 61 2 95656161

Email: n.shackel@centenary.usyd.edu.au
Seq primer: M13 forward.

FEATURES

source

Location/Qualifiers
1. 570
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="SSH library from Adult Cirrhotic Liver"
/note="Organ: Liver; Vector: pGEM T Easy II (Promega)"

ORIGIN

Query Match 19.8%; Score 168; DB 12; Length 570;
Best Local Similarity 99.4%; Pred. No. 2.2e-13;
Matches 179; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 ATGTGAGCAAGGGCGAGAGCTTTACCGGGGTGTGCCATCTCTGTGAGCTGAC 60
DB 391 ATGTGAGCAAGGGCGAGAGCTTTACCGGGGTGTGCCATCTCTGTGAGCTGAC 450
QY 61 GGGCAGCTAAAGGGCGCAAGTTGAGCTGTCCGGCGAGGGCGAGTGGCACTTAC 120
DB 451 GGGCAGCTAAAGGGCGCAAGTTGAGCTGTCCGGCGAGGGCGAGTGGCACTTAC 510
QY 121 GGCAGAGCTGACCCCTGAAGTTGATCTGACACCAACCGGC-AAAGCTGCCGTGCCCTGAGCCAC 179
DB 511 GGCAGAGCTGACCCCTGAAGTTGATCTGACACCAACCGGC-AAAGCTGCCGTGCCCTGAGCCAC 570

RESULT 14
BM888204 240 bp mRNA linear EST 08-MAR-2002
LOCUS TMW102 Human Trabecular Meshwork cDNA library Homo sapiens cDNA 5',
DEFINITION mRNA sequence.

ACCESSION BM888204
VERSION BM888204.1 GI:19271948
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE

1 (bases 1 to 240)
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS

Wirtz M.K., Samples J.R., Xu H., Severson T. and Acott T.S.
Expression Profile and Genome Location of cDNA Clones from an
Infant Human Trabecular Meshwork Library

TITLE

Unpublished (2002)

JOURNAL

COMMENT

Contact: Wirtz MK
Glaucoma Genetics Lab
Oregon Health Sciences University
3375 S.W. Terwilliger Blvd., Portland, OR 97201-4197, USA
Tel: 503-494-4698
Fax: 503-494-6875
Email: wirtzm@ohsu.edu

FEATURES

source

Location/Qualifiers
1. 240
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="Human Trabecular Meshwork"
/dev_stage="2 week to 2 year old infants"
/lab_host="TOP10"
/note="Vector: pCDNA3; Site 1: EcoRI; Site 2: EcoRI; Human
cDNA library made from mRNA isolated from 6 individuals,
meshwork cells established from eyes from 6 individuals,
ages 2 weeks to 2 years. Cells were harvested at passages
3 through 6. Invitrogen made a unidirectional cDNA library
from the mRNA from the frozen cells using a pCDNA3 vector
and TPO10F host cells."

ORIGIN

Query Match

19.0%; Score 161.8; DB 12; Length 240;

Best Local Similarity 94.6%; Pred. No. 1.3e-12;
Matches 174; Conservative 0; Mismatches 9; Indels 1; Gaps 1;

QY 1 ATGTGAGCAAGGGCGAGAGCTTTACCGGGGTGTGCCATCTCTGTGAGCTGAC 60
DB 57 ATGTGAGCAAGGGCGAGAGCTTTACCGGGGTGTGCCATCTCTGTGAGCTGAC 116
QY 61 GGGCAGCTAAAGGGCGCAAGTTGAGCTGTCCGGCGAGGGCGAGTGGCACTTAC 120
DB 117 GGGCAGCTAAAGGGCGCAAGTTGAGCTGTCCGGCGAGGGCGAGTGGCACTTAC 175
QY 121 GGCAGAGCTGACCCCTGAAGTTGATCTGACACCAACCGGC-AAAGCTGCCGTGCCCTGAGCCAC 180
DB 176 GGCAGAGCTGACCCCTGAAGTTGATCTGACACCAACCGGC-AAAGCTGCCGTGCCCTGAGCCAC 225
QY 181 CTCG 184
DB 236 CTAG 239

RESULT 15

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1, rue Laurent Fries, 67404 Illkirch Cedex, France
Tel: +33388653318
Fax: +33388653201
Email: dickmeis@cit.ust.fr
subtracted probe clones, enriched for nodal targets.
Location/Qualifiers
1. 525
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="TDSUBC 2D10"
/issue_type="whole embryo"
/dev_stage="shield stage, 6 hrs post-fertilisation"
/lab_host="TOP10 OneShot cells"
/clone_lib="Zebrafish shield stage whole embryo"
/note="Vector: pCRIT-TOPO, directly subcloned fragments
resulting from the nested PCR of the Clontech PERSselect
subtraction procedure"

ORIGIN

Query Match

Best Local Similarity

Matches

Conservative

Mismatches

Indels

Gaps

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

Db	122	GTCTTTACGACACCATTA	CTGTCACACAACTG	CGCCCTTCGAAAGATCC	CAAC	181
Qy	640	GAGAGCCGCGATCATG	GTCTGCTGAGTTCGTG	ACCGCCGCGGATCAT	CTCGGC	699
Db	182	GAAAAGAGAGACCAT	GTGTCCTTGAGTTG	TACAGCTGTGGGAT	TACACATG	241
Qy	700	ATGAGAGCTGTACAG	AGCTTAG	725		
Db	242	ATGATGAATATACAA	TATATCTAG	267		

Search completed: May 15, 2004, 05:20:30
 Job time : 2821 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2004, 02:57:45 ; Search time 88 Seconds
(without alignments)
5360.321 Million cell updates/sec

Title: US-09-931-232-2
Perfect score: 850
Sequence: 1 atcgatgacgaagcgagga.....tagatcatgttagatgc 850

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA:*

- 1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
- 2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
- 3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
- 4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
- 5: /cgn2_6/ptodata/2/ina/PTCTUS.COMB.seq:*
- 6: /cgn2_6/ptodata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	850	100.0	US-09-062-102-2	Sequence 2, Appli
2	850	100.0	US-09-364-946-2	Sequence 2, Appli
3	718.6	84.5	US-09-770-315-1	Sequence 1, Appli
4	718	84.5	US-09-503-799-1	Sequence 1, Appli
5	717.4	84.4	US-09-172-063-11	Sequence 11, Appli
6	717.4	84.4	US-09-316-919-12	Sequence 12, Appli
7	717.4	84.4	US-09-602-641-11	Sequence 11, Appli
8	717.4	84.4	US-09-920-922-1	Sequence 1, Appli
9	717.4	84.4	US-09-172-063-29	Sequence 29, Appli
10	717.4	84.4	US-09-602-641-29	Sequence 29, Appli
11	717.4	84.4	US-09-085-305-5	Sequence 5, Appli
12	717.4	84.4	US-09-417-197-128	Sequence 128, App
13	717.4	84.4	US-09-417-197-112	Sequence 112, App
14	717.4	84.4	US-09-417-197-58	Sequence 58, Appli
15	717.4	84.4	US-09-417-197-64	Sequence 64, Appli
16	717.4	84.4	US-09-417-197-62	Sequence 62, Appli
17	717.4	84.4	US-08-818-253-1	Sequence 1, Appli
18	717.4	84.4	US-08-818-253-5	Sequence 5, Appli
19	717.4	84.4	US-08-818-252-1	Sequence 1, Appli
20	717.4	84.4	US-08-818-252-5	Sequence 5, Appli
21	717.4	84.4	US-09-800-170-47	Sequence 47, Appli
22	717.4	84.4	US-09-417-197-74	Sequence 74, Appli
23	717.4	84.4	US-09-417-197-70	Sequence 70, Appli
24	717.4	84.4	US-09-417-197-140	Sequence 140, App
25	717.4	84.4	US-09-417-197-76	Sequence 76, Appli
26	717.4	84.4	US-09-417-197-60	Sequence 60, Appli
27	717.4	84.4	US-09-417-197-118	Sequence 118, App

28	717.4	84.4	US-09-591-025-8	Sequence 8, Appli
29	717.4	84.4	US-09-894-927B-8	Sequence 8, Appli
30	717.4	84.4	US-09-417-197-110	Sequence 110, App
31	717.4	84.4	US-09-417-197-136	Sequence 136, App
32	717.4	84.4	US-09-417-197-66	Sequence 66, Appli
33	717.4	84.4	US-09-417-197-122	Sequence 122, App
34	717.4	84.4	US-09-417-197-78	Sequence 78, Appli
35	717.4	84.4	US-09-417-197-132	Sequence 132, App
36	717.4	84.4	US-09-796-575-4	Sequence 4, Appli
37	717.4	84.4	US-09-796-575-5	Sequence 5, Appli
38	717.4	84.4	US-09-208-827-5	Sequence 5, Appli
39	717.4	84.4	US-10-043-074-5	Sequence 5, Appli
40	717.4	84.4	US-09-133-944-2	Sequence 2, Appli
41	717.4	84.4	US-09-208-827-2	Sequence 2, Appli
42	717.4	84.4	US-10-043-074-2	Sequence 2, Appli
43	717.4	84.4	US-09-133-944-1	Sequence 1, Appli
44	717.4	84.4	US-09-208-827-1	Sequence 1, Appli
45	717.4	84.4	US-10-043-074-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-062-102-2
Sequence 2, Application US/09062102
Patent No. 6130313
GENERAL INFORMATION:
APPLICANT: Li, Xiangfang
TITLE OF INVENTION: Rapidly Degrading GFP-Fusion Proteins and Methods
FILE REFERENCE: D6100
CURRENT APPLICATION NUMBER: US/09/062,102
CURRENT FILING DATE: 1998-04-17
EARLIER APPLICATION NUMBER: US 60/060, 855
EARLIER FILING DATE: 1997-10-02
NUMBER OF SEQ ID NOS: 3
SEQ ID NO 2
LENGTH: 850
TYPE: DNA
ORGANISM: artificial sequence
FEATURE:
OTHER INFORMATION: DNA sequence which encodes the EGFP-MODC422-461
OTHER INFORMATION: fusion protein.
US-09-062-102-2

Query Match 100.0%; Score 850; DB 3; Length 850;
Best Local Similarity 100.0%; Pred. No. 1.2e-160; Indels 0; Gaps 0;
Matches 850; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCTGACGAAGGCGAGAGCTGTTCAACGGGGTGGTGCCTGCTGAGCTGAGC 60
Db 1 ATGCTGACGAAGGCGAGAGCTGTTCAACGGGGTGGTGCCTGCTGAGCTGAGC 60
QY 61 GCGACGTAACCGCCCAAGTTCAGGCTCCGCGCGAGGCGGAGGCGGAGTCCACTAC 120
Db 61 GCGACGTAACCGCCCAAGTTCAGGCTCCGCGCGAGGCGGAGGCGGAGTCCACTAC 120
QY 121 GCGAGTGAACCGCGAGTTCATCTGACACCGGCAAGCTGCGCCCTGAGCCAGC 180
Db 121 GCGAGTGAACCGCGAGTTCATCTGACACCGGCAAGCTGCGCCCTGAGCCAGC 180
QY 181 CTCGTGACCAACCTGACCTAGCGGTCAGTTCAGCCGCTACCCGACCATAGAG 240
Db 181 CTCGTGACCAACCTGACCTAGCGGTCAGTTCAGCCGCTACCCGACCATAGAG 240
QY 241 CAGACGACCTTCTTCAAGTCCGATGCGGAGAGCTTCCGAGAGCGGAGCATCTTC 300
Db 241 CAGACGACCTTCTTCAAGTCCGATGCGGAGAGCTTCCGAGAGCGGAGCATCTTC 300
QY 301 TTCAGGACGAGGCGACCTCAAGACCGGCGGAGGTTGAAGTTCAGAGGCGAGACCTG 360
Db 301 TTCAGGACGAGGCGACCTCAAGACCGGCGGAGGTTGAAGTTCAGAGGCGAGACCTG 360

Db 301 TTCAAGGACGAGGCACTAACAAGCCGCGCGAGTGAAGTTCAGAGGCGACACCTTG 360
Qy 361 GTGAACCGCATCGAGCTGAGAGGCGATCATCTTCAAGAGAGCGGCAATCTGCGGAC 420
Db 361 GTGAACCGCATCGAGCTGAGAGGCGATCATCTTCAAGAGAGCGGCAATCTGCGGAC 420
Qy 421 AAGCTGAGTCAACTAACAAGCGCAACCGTCTATCATGCGCGAACAAGAAAC 480
Db 421 AAGCTGAGTCAACTAACAAGCGCAACCGTCTATCATGCGCGAACAAGAAAC 480
Qy 481 GGCACTCAAGTGAAGCTTCAAGATCCGCAACAATCGAGAGCGGAGCTGCACTGCGC 540
Db 481 GGCACTCAAGTGAAGCTTCAAGATCCGCAACAATCGAGAGCGGAGCTGCACTGCGC 540
Qy 541 GACCACTACGAGCAAGACCCCGCATCGCGCGAGCGCGCGCTGCTGCGCGAACAAC 600
Db 541 GACCACTACGAGCAAGACCCCGCATCGCGCGAGCGCGCGCTGCTGCGCGAACAAC 600
Qy 601 TACCTGAGACCCAGTCCGCGCGCTGAGCAAGACCCCAAGAGAGCGGATCATGCTG 660
Db 601 TACCTGAGACCCAGTCCGCGCGCTGAGCAAGACCCCAAGAGAGCGGATCATGCTG 660
Qy 661 CTGCTGAGTTCGTGACCGCGCGCGAGTCACTCTCGGATGAGCGAGCTGTAACAAG 720
Db 661 CTGCTGAGTTCGTGACCGCGCGCGAGTCACTCTCGGATGAGCGAGCTGTAACAAG 720
Qy 721 CTGAGCATGCTTCCCGCGCGAGTGAAGAGAGAGAGATGAGAGCGCTGCCATGCT 780
Db 721 CTGAGCATGCTTCCCGCGCGAGTGAAGAGAGAGAGATGAGAGCGCTGCCATGCT 780
Qy 781 TGTGCGCAGAGAGAGGAGGATGAGAGCGTCACTGAGCTGCTTCTGTAAGATCAAT 840
Db 781 TGTGCGCAGAGAGAGGAGGATGAGAGCGTCACTGAGCTGCTTCTGTAAGATCAAT 840
Qy 841 GTGTAGATGC 850
Db 841 GTGTAGATGC 850

RESULT 2
US-09-364-946-2
; Sequence 2, Application US/09364946
; Patent No. 6306600
; GENERAL INFORMATION:
; APPLICANT: Kain, Steve
; APPLICANT: Li, Xiangyang
; TITLE OF INVENTION: Rapidly Degrading GFP-Fusion Proteins and Methods
; FILE REFERENCE: D6100CIP/D2
; CURRENT APPLICATION NUMBER: US/09/364,946
; EARLIER FILING DATE: 1999-07-30
; EARLIER APPLICATION NUMBER: US 09/191,233
; EARLIER FILING DATE: 1998-11-13
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 2
; LENGTH: 850
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURES:
; OTHER INFORMATION: DNA sequence which encodes the EGFP-MODC422-461
; OTHER INFORMATION: fusion protein.
US-09-364-946-2

Query Match 100.0%; Score 850; DB 4; Length 850;
Best Local Similarity 100.0%; Pred. No. 1.2e-160; Indels 0; Gaps 0;
Matches 850; Conservative 0; Mismatches 0;

Qy 1 ATGGTGAAGAGGCGAGAGAGCTGTCACCGGGTGTGCTCCATCTGCTGAGCTGAGC 60
Db 1 ATGGTGAAGAGGCGAGAGAGCTGTCACCGGGTGTGCTCCATCTGCTGAGCTGAGC 60
Qy 61 GCGACGTAACCGGCAAGTTCAAGCTGTCGCGCGAGGCGAGGCGAGGATCCACTTAC 120
Db 61 GCGACGTAACCGGCAAGTTCAAGCTGTCGCGCGAGGCGAGGCGAGGATCCACTTAC 120

Db 61 GCGACGTAACCGGCAAGTTCAAGCTGTCGCGCGAGGCGAGGCGAGGATCCACTTAC 120
Qy 121 GGCAAGCTGACCCCTGAGATTCATCTGACCAAGGCAAGCTGCTGCTGCTGCTGCTGCT 180
Db 121 GGCAAGCTGACCCCTGAGATTCATCTGACCAAGGCAAGCTGCTGCTGCTGCTGCTGCT 180
Qy 181 CTGCTGACCAAGCTGACCTTCAAGTCCGCGATGCGAGAGGCTGCTGCTGCTGCTGCTGCT 240
Db 181 CTGCTGACCAAGCTGACCTTCAAGTCCGCGATGCGAGAGGCTGCTGCTGCTGCTGCTGCT 240
Qy 241 CAGCAGACCTTCAAGTCCGCGATGCGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
Db 241 CAGCAGACCTTCAAGTCCGCGATGCGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
Qy 301 TTCAAGGACGAGCGCAACTAACAAGCCGCGCGAGGCTGAAAGTTTGAAGGCGACCTTG 360
Db 301 TTCAAGGACGAGCGCAACTAACAAGCCGCGCGAGGCTGAAAGTTTGAAGGCGACCTTG 360
Qy 361 GTGAACCGCATCGAGCTGAGAGGCGATGATCTTCAAGAGAGAGCGGCAATCTGCGGAC 420
Db 361 GTGAACCGCATCGAGCTGAGAGGCGATGATCTTCAAGAGAGAGCGGCAATCTGCGGAC 420
Qy 421 AAGCTGAGTCAACTAACAAGCGCAACAAGCTCTATCATGCGCGAACAAGAAAC 480
Db 421 AAGCTGAGTCAACTAACAAGCGCAACAAGCTCTATCATGCGCGAACAAGAAAC 480
Qy 481 GGCACTCAAGTGAAGCTTCAAGATCCGCAACAATCGAGAGCGGAGCTGCACTGCGC 540
Db 481 GGCACTCAAGTGAAGCTTCAAGATCCGCAACAATCGAGAGCGGAGCTGCACTGCGC 540
Qy 541 GACCACTACGAGCAAGACCCCGCATCGCGCGAGCTGCTGCGCGAACAAC 600
Db 541 GACCACTACGAGCAAGACCCCGCATCGCGCGAGCTGCTGCGCGAACAAC 600
Qy 601 TACCTGAGACCCAGTCCGCGCGCTGAGCAAGACCCCAAGAGAGCGGATCATGCTG 660
Db 601 TACCTGAGACCCAGTCCGCGCGCTGAGCAAGACCCCAAGAGAGCGGATCATGCTG 660
Qy 661 CTGCTGAGTTCGTGACCGCGCGCGAGTCACTCTCGGATGAGCGAGCTGTAACAAG 720
Db 661 CTGCTGAGTTCGTGACCGCGCGCGAGTCACTCTCGGATGAGCGAGCTGTAACAAG 720
Qy 721 CTGAGCATGCTTCCCGCGCGAGTGAAGAGAGAGATGAGAGCGCTGCCATGCT 780
Db 721 CTGAGCATGCTTCCCGCGCGAGTGAAGAGAGAGATGAGAGCGCTGCCATGCT 780
Qy 781 TGTGCGCAGAGAGAGGAGGATGAGAGCGTCACTGAGCTGCTTCTGTAAGATCAAT 840
Db 781 TGTGCGCAGAGAGAGGAGGATGAGAGCGTCACTGAGCTGCTTCTGTAAGATCAAT 840
Qy 841 GTGTAGATGC 850
Db 841 GTGTAGATGC 850

RESULT 3
US-09-770-315-1
; Sequence 1, Application US/09770315
; Patent No. 6429001
; GENERAL INFORMATION:
; APPLICANT: Chiron Corporation
; TITLE OF INVENTION: Recombinant AAV Packaging Systems
; FILE REFERENCE: 20263-501
; CURRENT APPLICATION NUMBER: US/09/770,315
; EARLIER FILING DATE: 2001-01-26
; EARLIER APPLICATION NUMBER: US 60/178,536
; EARLIER FILING DATE: 2000-01-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 7015
; TYPE: DNA
; ORGANISM: Unknown

FEATURE:
OTHER INFORMATION: recombinant DNA
US-09-770-315-1

Query Match
Best Local Similarity 84.5%; Score 718.6; DB 4; Length 7015;
Pred. No. 2.1e-134;
Matches 724; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

```
QY 1 ATGGTGAAGCAAGGCGGAGAGCTTTACCGGGGTGTGCCCATCTGTGAGCTGAGC 60
DB ATGGTGAAGCAAGGCGGAGAGCTTTACCGGGGTGTGCCCATCTGTGAGCTGAGC 60
QY 3171 ATGGTGAAGCAAGGCGGAGAGCTTTACCGGGGTGTGCCCATCTGTGAGCTGAGC 3230
DB 3171 ATGGTGAAGCAAGGCGGAGAGCTTTACCGGGGTGTGCCCATCTGTGAGCTGAGC 3230
QY 61 GGGACGCTAAACGGCCCAAGTTCAAGCTGCTCGGCGAAGGCGAGGCGCATGCCACTAC 120
DB 3231 GGGACGCTAAACGGCCCAAGTTCAAGCTGCTCGGCGAAGGCGAGGCGCATGCCACTAC 120
QY 121 GGGACGCTAAACGGCCCAAGTTCAAGCTGCTCGGCGAAGGCGAGGCGCATGCCACTAC 3290
DB 3291 GGGACGCTAAACGGCCCAAGTTCAAGCTGCTCGGCGAAGGCGAGGCGCATGCCACTAC 3290
QY 181 CTGGTGAACCACTTGAAGCTTCAAGCTGCTCGGCGAAGGCGAGGCGCATGCCACTAC 240
DB 3351 CTGGTGAACCACTTGAAGCTTCAAGCTGCTCGGCGAAGGCGAGGCGCATGCCACTAC 240
QY 241 CAGCAAGCACTTCTTCAAGCTTCAAGCTGCTCGGCGAAGGCGAGGCGCATGCCACTAC 300
DB 3411 CAGCAAGCACTTCTTCAAGCTTCAAGCTGCTCGGCGAAGGCGAGGCGCATGCCACTAC 300
QY 301 TTCAAGGAGAGCGGCACTAAGAACCCGCGCGAGGCTGAAGTTGAGGGGCAACCCCTG 3470
DB 3471 TTCAAGGAGAGCGGCACTAAGAACCCGCGCGAGGCTGAAGTTGAGGGGCAACCCCTG 3470
QY 361 GTGAACCGGATGAGAGCTGAAGGCGATTCGATGAAGAGAGCGGCAATCTTGGGGCAC 3530
DB 3531 GTGAACCGGATGAGAGCTGAAGGCGATTCGATGAAGAGAGCGGCAATCTTGGGGCAC 3530
QY 421 AAGCTGAGTAACTAACAAGCAAGCAACGCTTAATCATAGGCGGCAAGAGAGAGAC 480
DB 3591 AAGCTGAGTAACTAACAAGCAAGCAACGCTTAATCATAGGCGGCAAGAGAGAGAC 480
QY 481 GGCATCAAGGTGAAGTTGAAGTCCGCCAACAATGAGAGAGCGGAGGTGAGTCCGCC 540
DB 3651 GGCATCAAGGTGAAGTTGAAGTCCGCCAACAATGAGAGAGCGGAGGTGAGTCCGCC 540
QY 541 GACCACTACGAGAGAACCCCATTCGCGAGCGGCCCGTGTGCTGCTCCGACACAC 600
DB 3711 GACCACTACGAGAGAACCCCATTCGCGAGCGGCCCGTGTGCTGCTCCGACACAC 600
QY 601 TACCTGAGCAACCGAGTCCGCTGAGCAAGAACCCCAAGAGCGGATCAATGATGTC 660
DB 3771 TACCTGAGCAACCGAGTCCGCTGAGCAAGAACCCCAAGAGCGGATCAATGATGTC 660
QY 661 CTGCTGAGATTGCTGACCGCGCGCGGAGTCACTCTCGGCAATGAGAGCTGTACAAGAG 720
DB 3831 CTGCTGAGATTGCTGACCGCGCGCGGAGTCACTCTCGGCAATGAGAGCTGTACAAGAG 720
QY 721 CTGCTGAGATTGCTGACCGCGCGCGGAGTCACTCTCGGCAATGAGAGCTGTACAAGAG 3890
DB 3891 AGGCGCGCGCACT 3903
```

RESULT 4
US-09-503-799-1
Sequence 1, Application US/09503799
Patent No. 6451563

GENERAL INFORMATION:
APPLICANT: Willems, Burghardt
APPLICANT: Junghans, Claas
TITLE OF INVENTION: METHOD FOR MAKING LINEAR, COVALENTLY CLOSED DNA CONSTRUCTS
FILE REFERENCE: NHL-NP-21
CURRENT APPLICATION NUMBER: US/09/503, 799
CURRENT FILING DATE: 2000-02-14

NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 4397
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: pUC 19 derivative, comprising a coding sequence for the modified
OTHER INFORMATION: enhanced green fluorescence protein from *Aequorea victoria* under
US-09-503-799-1

Query Match
Best Local Similarity 100.0%; Score 718; DB 4; Length 4397;
Pred. No. 2.6e-134;
Matches 718; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGGTGAAGCAAGGCGGAGAGCTTTACCGGGGTGTGCCCATCTGTGAGCTGAGC 60
DB 3073 ATGGTGAAGCAAGGCGGAGAGCTTTACCGGGGTGTGCCCATCTGTGAGCTGAGC 3132
QY 61 GGGACGCTAAACGGCCCAAGTTCAAGCTGCTCGGCGAAGGCGAGGCGCATGCCACTAC 120
DB 3133 GGGACGCTAAACGGCCCAAGTTCAAGCTGCTCGGCGAAGGCGAGGCGCATGCCACTAC 120
QY 121 GGGACGCTAAACGGCCCAAGTTCAAGCTGCTCGGCGAAGGCGAGGCGCATGCCACTAC 3192
DB 3193 GGGACGCTAAACGGCCCAAGTTCAAGCTGCTCGGCGAAGGCGAGGCGCATGCCACTAC 3192
QY 181 CTGGTGAACCACTTGAAGCTTCAAGCTGCTCGGCGAAGGCGAGGCGCATGCCACTAC 240
DB 3253 CTGGTGAACCACTTGAAGCTTCAAGCTGCTCGGCGAAGGCGAGGCGCATGCCACTAC 240
QY 241 CAGCAAGCACTTCTTCAAGCTTCAAGCTGCTCGGCGAAGGCGAGGCGCATGCCACTAC 300
DB 3313 CAGCAAGCACTTCTTCAAGCTTCAAGCTGCTCGGCGAAGGCGAGGCGCATGCCACTAC 300
QY 301 TTCAAGGAGAGCGGCACTAAGAACCCGCGCGAGGCTGAAGTTGAGGGGCAACCCCTG 3372
DB 3373 TTCAAGGAGAGCGGCACTAAGAACCCGCGCGAGGCTGAAGTTGAGGGGCAACCCCTG 3372
QY 361 GTGAACCGGATGAGAGCTGAAGGCGATTCGATGAAGAGAGCGGCAATCTTGGGGCAC 420
DB 3433 GTGAACCGGATGAGAGCTGAAGGCGATTCGATGAAGAGAGCGGCAATCTTGGGGCAC 420
QY 421 AAGCTGAGTAACTAACAAGCAAGCAACGCTTAATCATAGGCGGCAAGAGAGAGAC 480
DB 3493 AAGCTGAGTAACTAACAAGCAAGCAACGCTTAATCATAGGCGGCAAGAGAGAGAC 480
QY 481 GGCATCAAGGTGAAGTTGAAGTCCGCCAACAATGAGAGAGCGGAGGTGAGTCCGCC 540
DB 3553 GGCATCAAGGTGAAGTTGAAGTCCGCCAACAATGAGAGAGCGGAGGTGAGTCCGCC 540
QY 541 GACCACTACGAGAGAACCCCATTCGCGAGCGGCCCGTGTGCTGCTCCGACACAC 600
DB 3613 GACCACTACGAGAGAACCCCATTCGCGAGCGGCCCGTGTGCTGCTCCGACACAC 600
QY 601 TACCTGAGCAACCGAGTCCGCTGAGCAAGAACCCCAAGAGCGGATCAATGATGTC 660
DB 3673 TACCTGAGCAACCGAGTCCGCTGAGCAAGAACCCCAAGAGCGGATCAATGATGTC 660
QY 661 CTGCTGAGATTGCTGACCGCGCGCGGAGTCACTCTCGGCAATGAGAGCTGTACAAGAG 718
DB 3733 CTGCTGAGATTGCTGACCGCGCGCGGAGTCACTCTCGGCAATGAGAGCTGTACAAGAG 718
```

RESULT 5
US-09-172-063-11

Sequence 11, Application US/09172063
Patent No. 6150176
GENERAL INFORMATION:
APPLICANT: Tsien, Roger Y.
APPLICANT: Miyawaki, Atsushi
APPLICANT: Llopis, Juan

APPLICANT: Machter, Rebekka M.
 APPLICANT: Remington, S. James
 TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
 FILE REFERENCE: 07257/071001
 CURRENT APPLICATION NUMBER: US/09/172,063
 EARLIER FILING DATE: 1998-10-13
 EARLIER APPLICATION NUMBER: 09/094,359
 NUMBER OF SEQ ID NOS: 38
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 11
 LENGTH: 720
 TYPE: DNA
 ORGANISM: Aequorea victoria
 FEATURE:
 NAME/KEY: misc. feature
 LOCATION: (0)..(0)
 OTHER INFORMATION: EGFP
 US-09-172-063-11

Query Match 84.4%; Score 717.4; DB 3; Length 720;
 Best Local Similarity 99.9%; Pred. No. 2.7e-134;
 Matches 718; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	ATGTTGAGCAAGGCGGAGAGCTTTCAACCGGGTGTGCTCCATCTGTGAGCTGAGC	60
DB	1	ATGTTGAGCAAGGCGGAGAGCTTTCAACCGGGTGTGCTCCATCTGTGAGCTGAGC	60
QY	61	GCGCAGTAAACGGCCACAAGTTCAAGCGTGTCCGGGAGAGGGGCGATGCCCTTAC	120
DB	61	GCGCAGTAAACGGCCACAAGTTCAAGCGTGTCCGGGAGAGGGGCGATGCCCTTAC	120
QY	121	GCGAAGCTGACCCCTGAAGTTCACTGACACCGGCAAGCTGCCGTGCTGCGCCACC	180
DB	121	GCGAAGCTGACCCCTGAAGTTCACTGACACCGGCAAGCTGCCGTGCTGCGCCACC	180
QY	181	CTGTGACCAACCTTCAAGTCCGCGATGAGCTTCAAGCGGCAAGCTGCCGTGCTGCGCCACC	240
DB	181	CTGTGACCAACCTTCAAGTCCGCGATGAGCTTCAAGCGGCAAGCTGCCGTGCTGCGCCACC	240
QY	241	CAGCAGCACTTCTTCAAGTCCGCGATGAGCTTCAAGCGGCAAGCTGCCGTGCTGCGCCACC	300
DB	241	CAGCAGCACTTCTTCAAGTCCGCGATGAGCTTCAAGCGGCAAGCTGCCGTGCTGCGCCACC	300
QY	301	TTCAAGGACGAGCGGCACTTCAAGCGGCAAGCTGCCGTGCTGCGCCACC	360
DB	301	TTCAAGGACGAGCGGCACTTCAAGCGGCAAGCTGCCGTGCTGCGCCACC	360
QY	361	GTGAACCGCATGAGCTGAGGAGGAGCTTCAAGCGGCAAGCTGCCGTGCTGCGCCACC	420
DB	361	GTGAACCGCATGAGCTGAGGAGGAGCTTCAAGCGGCAAGCTGCCGTGCTGCGCCACC	420
QY	421	AAAGTGAAGTCACTTCAAGCGGCAAGCTGCCGTGCTGCGCCACC	480
DB	421	AAAGTGAAGTCACTTCAAGCGGCAAGCTGCCGTGCTGCGCCACC	480
QY	481	GCGATCAAGGTGAAGTTCAAGCGGCAAGCTGCCGTGCTGCGCCACC	540
DB	481	GCGATCAAGGTGAAGTTCAAGCGGCAAGCTGCCGTGCTGCGCCACC	540
QY	541	GACCACTTCAAGCGGCAAGCTGCCGTGCTGCGCCACC	600
DB	541	GACCACTTCAAGCGGCAAGCTGCCGTGCTGCGCCACC	600
QY	601	TACCTGAGCAACCGGCTGAGGAGGAGCTTCAAGCGGCAAGCTGCCGTGCTGCGCCACC	660
DB	601	TACCTGAGCAACCGGCTGAGGAGGAGCTTCAAGCGGCAAGCTGCCGTGCTGCGCCACC	660
QY	661	CTGCTGAGGTTGTGACCGCGCGCGGAGTCACTCTGCGGATGAGAGAGCTGTACAAGAA	719
DB	661	CTGCTGAGGTTGTGACCGCGCGCGGAGTCACTCTGCGGATGAGAGAGCTGTACAAGAA	719

RESULT 6
 US-09-316-919-12
 Sequence 12, Application US/09316919
 Patent No. 6469154
 GENERAL INFORMATION:
 APPLICANT: Telen, Roger Y.
 APPLICANT: Baird, Geoffrey
 TITLE OF INVENTION: FLUORESCENT PROTEIN INDICATORS
 FILE REFERENCE: 07257/073001
 CURRENT APPLICATION NUMBER: US/09/316,919
 NUMBER OF SEQ ID NOS: 63
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 12
 LENGTH: 720
 TYPE: DNA
 ORGANISM: Aequorea victoria
 US-09-316-919-12

Query Match 84.4%; Score 717.4; DB 4; Length 720;
 Best Local Similarity 99.9%; Pred. No. 2.7e-134;
 Matches 718; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	ATGTTGAGCAAGGCGGAGAGCTTTCAACCGGGTGTGCTCCATCTGTGAGCTGAGC	60
DB	1	ATGTTGAGCAAGGCGGAGAGCTTTCAACCGGGTGTGCTCCATCTGTGAGCTGAGC	60
QY	61	GCGCAGTAAACGGCCACAAGTTCAAGCGTGTCCGGGAGAGGGGCGATGCCCTTAC	120
DB	61	GCGCAGTAAACGGCCACAAGTTCAAGCGTGTCCGGGAGAGGGGCGATGCCCTTAC	120
QY	121	GCGAAGCTGACCCCTGAAGTTCACTGACACCGGCAAGCTGCCGTGCTGCGCCACC	180
DB	121	GCGAAGCTGACCCCTGAAGTTCACTGACACCGGCAAGCTGCCGTGCTGCGCCACC	180
QY	181	CTGTGACCAACCTTCAAGTCCGCGATGAGCTTCAAGCGGCAAGCTGCCGTGCTGCGCCACC	240
DB	181	CTGTGACCAACCTTCAAGTCCGCGATGAGCTTCAAGCGGCAAGCTGCCGTGCTGCGCCACC	240
QY	241	CAGCAGCACTTCTTCAAGTCCGCGATGAGCTTCAAGCGGCAAGCTGCCGTGCTGCGCCACC	300
DB	241	CAGCAGCACTTCTTCAAGTCCGCGATGAGCTTCAAGCGGCAAGCTGCCGTGCTGCGCCACC	300
QY	301	TTCAAGGACGAGCGGCACTTCAAGCGGCAAGCTGCCGTGCTGCGCCACC	360
DB	301	TTCAAGGACGAGCGGCACTTCAAGCGGCAAGCTGCCGTGCTGCGCCACC	360
QY	361	GTGAACCGCATGAGCTGAGGAGGAGCTTCAAGCGGCAAGCTGCCGTGCTGCGCCACC	420
DB	361	GTGAACCGCATGAGCTGAGGAGGAGCTTCAAGCGGCAAGCTGCCGTGCTGCGCCACC	420
QY	421	AAAGTGAAGTCACTTCAAGCGGCAAGCTGCCGTGCTGCGCCACC	480
DB	421	AAAGTGAAGTCACTTCAAGCGGCAAGCTGCCGTGCTGCGCCACC	480
QY	481	GCGATCAAGGTGAAGTTCAAGCGGCAAGCTGCCGTGCTGCGCCACC	540
DB	481	GCGATCAAGGTGAAGTTCAAGCGGCAAGCTGCCGTGCTGCGCCACC	540
QY	541	GACCACTTCAAGCGGCAAGCTGCCGTGCTGCGCCACC	600
DB	541	GACCACTTCAAGCGGCAAGCTGCCGTGCTGCGCCACC	600
QY	601	TACCTGAGCAACCGGCTGAGGAGGAGCTTCAAGCGGCAAGCTGCCGTGCTGCGCCACC	660
DB	601	TACCTGAGCAACCGGCTGAGGAGGAGCTTCAAGCGGCAAGCTGCCGTGCTGCGCCACC	660
QY	661	CTGCTGAGGTTGTGACCGCGCGCGGAGTCACTCTGCGGATGAGAGAGCTGTACAAGAA	719
DB	661	CTGCTGAGGTTGTGACCGCGCGCGGAGTCACTCTGCGGATGAGAGAGCTGTACAAGAA	719

Db 493 CAGCAGACTTCTTCAAGTCCGCCATGCCCCAGAGCTACGTCCAGAGCGACCATCTTC 552
Qy 301 TTCAAGAGCAGACGGCACTACAAAGACCCGCGCCGAGTGAAGTTGAGGGGACACACCTTG 360
Db 553 TTCAAGAGCAGACGGCACTACAAAGACCCGCGCCGAGTGAAGTTGAGGGGACACACCTTG 612
Qy 361 GTGAACCCGATCGAGCTGAAGGCGATCGACTTCAAGAGAGAGCGGCAACATCTCGGGGAC 420
Db 613 GTGAACCCGATCGAGCTGAAGGCGATCGACTTCAAGAGAGAGCGGCAACATCTCGGGGAC 672
Qy 421 AAGCTGAGTCACTACAAAGCAGCAGCAACGCTCTATCAAGGCGGCAACAGCAGAGAGAC 480
Db 673 AAGCTGAGTCACTACAAAGCAGCAGCAACGCTCTATCAAGGCGGCAACAGCAGAGAGAC 732
Qy 481 GGCATCAAGGTGAAGTTCAGATCCGCCAACAATCGAGAGCGGAGCTGAGCTGCGC 540
Db 733 GGCATCAAGGTGAAGTTCAGATCCGCCAACAATCGAGAGCGGAGCTGAGCTGCGC 792
Qy 541 GACCACTACAGCAGAGAACACCCCATCGGCGAGCGCCCGTGTGCTGCTCCCGACACAC 600
Db 793 GACCACTACAGCAGAGAACACCCCATCGGCGAGCGCCCGTGTGCTGCTCCCGACACAC 852
Qy 601 TACCTGAGCAGCAGCAGTCCGCTGAGCAAGACCCCAACAGAGCGGAGTCAATGCTC 660
Db 853 TACCTGAGCAGCAGCAGTCCGCTGAGCAAGACCCCAACAGAGCGGAGTCAATGCTC 912
Qy 661 CTGCTGAGTTCGTGACCGCGCGCGGATGCACTCTGCGCATGAGAGAGCTGTAAGAA 719
Db 913 CTGCTGAGTTCGTGACCGCGCGCGGATGCACTCTGCGCATGAGAGAGCTGTAAGAA 971

RESULT 11

US-09-085-305-5
Sequence 5, Application US/09085305
Patent No. 6191269
GENERAL INFORMATION:
APPLICANT: Pollock, Allan
APPLICANT: Lovett, David H.
TITLE OF INVENTION: Selective Induction of Apoptosis in
TITLE OF INVENTION: Malignant Cancer Cells by Delivery of N-Terminal
TITLE OF INVENTION: Interleukin-1-Alpha Pro-Piece Polypeptide
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESSES:
ADDRESSES: Bozicevic & Reed, LLP
STREET: 285 Hamilton Ave, Suite 200
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/085,305
FILING DATE: 29-MAY-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Francis, Carol L
REGISTRATION NUMBER: 36,513
REFERENCE/DOCKET NUMBER: 6510/102US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-327-3400
TELEFAX: 650-327-3231
TELEX:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1095 base pairs

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 1...1092
OTHER INFORMATION:
US-09-085-305-5

Query Match 84.4%; Score 717.4; DB 3; Length 1095;
Best Local Similarity 99.9%; Pred. No. 2.8e-134;
Matches 718; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGGTGAAGCAAGGGCGAGAGCTGTTCACCGGGGTGGTGGCCATCTGCTGAGCTGAGAC 60
Db 376 ATGGTGAAGCAAGGGCGAGAGCTGTTCACCGGGGTGGTGGCCATCTGCTGAGCTGAGAC 435
Qy 61 GCGACGTAACCGGCAAAAGTTCCAGGTGTCCGCGAGAGGCGAGAGGCGATGCACTAC 120
Db 436 GCGACGTAACCGGCAAAAGTTCCAGGTGTCCGCGAGAGGCGAGAGGCGATGCACTAC 495
Qy 121 GCGAAGTGAACCTGGAAGTTCTGTGACCAACCGGCAAGCTGCTGCTGCTGCTGCTGCTG 180
Db 496 GCGAAGTGAACCTGGAAGTTCTGTGACCAACCGGCAAGCTGCTGCTGCTGCTGCTGCTG 555
Qy 181 CTGCTGACCAACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTG 240
Db 556 CTGCTGACCAACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTG 615
Qy 241 CAGCAGACTTCTTCAAGTCCGCGATGCCGAAAGGCTACGTCAGAGCGCACCATCTTC 300
Db 616 CAGCAGACTTCTTCAAGTCCGCGATGCCGAAAGGCTACGTCAGAGCGCACCATCTTC 675
Qy 301 TTCAAGAGCAGACGGCACTCAAGACCCGCGGAGGTGAAGTTGAGGGGCAACCTTG 360
Db 676 TTCAAGAGCAGACGGCACTCAAGACCCGCGGAGGTGAAGTTGAGGGGCAACCTTG 735
Qy 361 GTGAACCCGATCGAGCTGAAGGCGATCGACTTCAAGAGAGAGCGGCAACATCTCGGGGAC 420
Db 736 GTGAACCCGATCGAGCTGAAGGCGATCGACTTCAAGAGAGAGCGGCAACATCTCGGGGAC 795
Qy 421 AAGCTGAGTCACTACAAAGCAGCAGCAACGCTCTATCAAGGCGGCAACAGCAGAGAGAC 480
Db 796 AAGCTGAGTCACTACAAAGCAGCAGCAACGCTCTATCAAGGCGGCAACAGCAGAGAGAC 855
Qy 481 GGCATCAAGGTGAAGTTCAGATCCGCCAACAATCGAGAGCGGAGCTGAGCTGCGC 540
Db 856 GGCATCAAGGTGAAGTTCAGATCCGCCAACAATCGAGAGCGGAGCTGAGCTGCGC 915
Qy 541 GACCACTACAGCAGAGAACACCCCATCGGCGAGCGCCCGTGTGCTGCTCCCGACACAC 600
Db 916 GACCACTACAGCAGAGAACACCCCATCGGCGAGCGCCCGTGTGCTGCTCCCGACACAC 975
Qy 601 TACCTGAGCAGCAGTCCGCTGAGCAAGACCCCAACAGAGGCGGATCAATGCTC 660
Db 976 TACCTGAGCAGCAGTCCGCTGAGCAAGACCCCAACAGAGGCGGATCAATGCTC 1035
Qy 661 CTGCTGAGTTCGTGACCGCGCGCGGATGCACTCTGCGCATGAGAGAGCTGTCAAGAA 719
Db 1036 CTGCTGAGTTCGTGACCGCGCGCGGATGCACTCTGCGCATGAGAGAGCTGTCAAGAA 1094

RESULT 12

US-09-417-197-128
Sequence 128, Application US/09417197
Patent No. 6518021
GENERAL INFORMATION:
APPLICANT: Ole Thastrup, et al.
TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An Ir
TITLE OF INVENTION: On A Cellular Response
FILE REFERENCE: 3759-0110P
CURRENT APPLICATION NUMBER: US/09/417,197

; CURRENT FILING DATE: 1999-10-07
 ; NUMBER OF SEQ ID NOS: 143
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 128
 ; LENGTH: 1140
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: actin-binding-domain-Egfp fusion
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(1137)
 ; US-09-417-197-128

Query Match 84.4%; Score 717.4; DB 4; Length 1140;
 Best Local Similarity 99.9%; Pred. No. 2.9e-134;
 Matches 718; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	ATGTGAGCAAGGCGGAGAGCTGTTACCGGGGTGTGCGCCATCTGTGAGCTGGAC	60
DB	421	ATGTGAGCAAGGCGGAGAGAGCTGTTACCGGGGTGTGCGCCATCTGTGAGCTGGAC	480
QY	61	GCGCAGCTAAACGGCCACAAGTTCAAGCTGTCGGGCGAGGGGAGTGGCCACTTAC	120
DB	481	GCGCAGCTAAACGGCCACAAGTTCAAGCTGTCGGGCGAGGGGAGTGGCCACTTAC	540
QY	121	GCGCAGCTGAGCCCTGAAGTTCAATCTGACACACCGGCAAGCTGCGCCCTGCGCCAC	180
DB	541	GCGCAGCTGAGCCCTGAAGTTCAATCTGACACACCGGCAAGCTGCGCCCTGCGCCAC	600
QY	181	CTCGTGAACACCCCTGACCTTACGAGCGGTGACGCTTACCGCCGTCACCCGACCAATGAAG	240
DB	601	CTCGTGAACACCCCTGACCTTACGAGCGGTGACGCTTACCGCCGTCACCCGACCAATGAAG	660
QY	241	CAGCAGCACTTCTTCAAGTCCGCGCCATGCGCCGAGGCTACGTCACGAGGCGACCACTTTC	300
DB	661	CAGCAGCACTTCTTCAAGTCCGCGCCATGCGCCGAGGCTACGTCACGAGGCGACCACTTTC	720
QY	301	TTCAAAGACGACGCGCACTTCAAGACCCGCGCGAGGTGAAGTTGAGGGCGACACCTTG	360
DB	721	TTCAAAGACGACGCGCACTTCAAGACCCGCGCGAGGTGAAGTTGAGGGCGACACCTTG	780
QY	361	GTGAACCGCATCGAGCTGAGGAGGCTGACCTTCAAGAGGACGCGAACAATCTCTGGGCGAC	420
DB	781	GTGAACCGCATCGAGCTGAGGAGGCTGACCTTCAAGAGGACGCGAACAATCTCTGGGCGAC	840
QY	421	AAGCTGAGTAACTACACAGCCCAACGCTTATATCATGCGCGACCAAGAGAAC	480
DB	841	AAGCTGAGTAACTACACAGCCCAACGCTTATATCATGCGCGACCAAGAGAAC	900
QY	481	GCGATCAAGGTGAATTTCAAGTCCGCGACCAATCGAGACGCGAGCGTGCAGCTGCGC	540
DB	901	GCGATCAAGGTGAATTTCAAGTCCGCGACCAATCGAGACGCGAGCGTGCAGCTGCGC	960
QY	541	GACCACTACAGCAAGAACACCCCATCGGCGACCGCGCGTGTGCGCGACCAACAC	600
DB	961	GACCACTACAGCAAGAACACCCCATCGGCGACCGCGCGTGTGCGCGACCAACAC	1020
QY	601	TACCTGAGCAACCAAGTCCGCGTGAAGAAACCCCAAGAGCGCGATCAATGTTC	660
DB	1021	TACCTGAGCAACCAAGTCCGCGTGAAGAAACCCCAAGAGCGCGATCAATGTTC	1080
QY	661	CTGCTGAGGTTCTGAGACCGCGCGCGGATCACTCTCGGATGAGAGAGCTGTAACAATA	719
DB	1081	CTGCTGAGGTTCTGAGACCGCGCGCGGATCACTCTCGGATGAGAGAGCTGTAACAATA	1139

RESULT 13
 US-09-417-197-112
 ; Sequence 112 Application US/09417197
 ; Patent No. 6518021
 ; GENERAL INFORMATION:
 ; APPLICANT: Ole THASTRUP, et al.

; TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An Ir
 ; FILE REFERENCE: 3759-0110P
 ; CURRENT APPLICATION NUMBER: US/09/417,197
 ; CURRENT FILING DATE: 1999-10-07
 ; NUMBER OF SEQ ID NOS: 143
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 112
 ; LENGTH: 1635
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: CDK2-Egfp fusion
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(1632)
 ; US-09-417-197-112

Query Match 84.4%; Score 717.4; DB 4; Length 1635;
 Best Local Similarity 99.9%; Pred. No. 3e-134;
 Matches 718; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	ATGTGAGCAAGGCGGAGAGCTGTTACCGGGGTGTGCGCCATCTGTGAGCTGGAC	60
DB	916	ATGTGAGCAAGGCGGAGAGAGCTGTTACCGGGGTGTGCGCCATCTGTGAGCTGGAC	975
QY	61	GCGCAGCTAAACGGCCACAAGTTCAAGCTGTCGGGCGAGGGGAGTGGCCACTTAC	120
DB	976	GCGCAGCTAAACGGCCACAAGTTCAAGCTGTCGGGCGAGGGGAGTGGCCACTTAC	1035
QY	121	GCGCAGCTGAGCCCTGAAGTTCAATCTGACACACCGGCAAGCTGCGCCCTGCGCCAC	180
DB	1036	GCGCAGCTGAGCCCTGAAGTTCAATCTGACACACCGGCAAGCTGCGCCCTGCGCCAC	1095
QY	181	CTCGTGAACACCCCTGACCTTACGAGCGGTGACGCTTACCGCCGTCACCCGACCAATGAAG	240
DB	1096	CTCGTGAACACCCCTGACCTTACGAGCGGTGACGCTTACCGCCGTCACCCGACCAATGAAG	1155
QY	241	CAGCAGCACTTCTTCAAGTCCGCGCCATGCGCCGAGGCTACGTCACGAGGCGACCACTTTC	300
DB	1156	CAGCAGCACTTCTTCAAGTCCGCGCCATGCGCCGAGGCTACGTCACGAGGCGACCACTTTC	1215
QY	301	TTCAAAGACGACGCGCACTTCAAGACCCGCGCGAGGTGAAGTTGAGGGCGACACCTTG	360
DB	1216	TTCAAAGACGACGCGCACTTCAAGACCCGCGCGAGGTGAAGTTGAGGGCGACACCTTG	1275
QY	361	GTGAACCGCATCGAGCTGAGGAGGCTGACCTTCAAGAGGACGCGAACAATCTCTGGGCGAC	420
DB	1276	GTGAACCGCATCGAGCTGAGGAGGCTGACCTTCAAGAGGACGCGAACAATCTCTGGGCGAC	1335
QY	421	AAGCTGAGTAACTACACAGCCCAACGCTTATATCATGCGCGACCAAGAGAAC	480
DB	1336	AAGCTGAGTAACTACACAGCCCAACGCTTATATCATGCGCGACCAAGAGAAC	1395
QY	481	GCGATCAAGGTGAATTTCAAGTCCGCGACCAATCGAGACGCGAGCGTGCAGCTGCGC	540
DB	1396	GCGATCAAGGTGAATTTCAAGTCCGCGACCAATCGAGACGCGAGCGTGCAGCTGCGC	1455
QY	541	GACCACTACAGCAAGAACACCCCATCGGCGACCGCGCGTGTGCGCGACCAACAC	600
DB	1456	GACCACTACAGCAAGAACACCCCATCGGCGACCGCGCGTGTGCGCGACCAACAC	1515
QY	601	TACCTGAGCAACCAAGTCCGCGTGAAGAAACCCCAAGAGCGCGATCAATGTTC	660
DB	1516	TACCTGAGCAACCAAGTCCGCGTGAAGAAACCCCAAGAGCGCGATCAATGTTC	1575
QY	661	CTGCTGAGGTTCTGAGACCGCGCGCGGATCACTCTCGGATGAGAGAGCTGTAACAATA	719
DB	1576	CTGCTGAGGTTCTGAGACCGCGCGCGGATCACTCTCGGATGAGAGAGCTGTAACAATA	1634

RESULT 14
 US-09-417-197-58

; Sequence 58, Application US/09417197
; Patent No. 6518021
; GENERAL INFORMATION:
; APPLICANT: Ole Thastrup, et al.
; TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An I
; TITLE OF INVENTION: On A Cellular Response
; FILE REFERENCE: 3759-0110P
; CURRENT APPLICATION NUMBER: US/09/417,197
; CURRENT FILING DATE: 1999-10-07
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 58
; LENGTH: 1815
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Ert2-BGFP fusion
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1812)
US-09-417-197-58

Query Match 84.4%; Score 717.4; DB 4; Length 1815;
Best Local Similarity 99.9%; Pred. No. 3e-134; Indels 0; Gaps 0;
Matches 718; Conservative 0; Mismatches 1;

QY 1 ATGTTAGCAAGGCGAGAGCTTTCAACGGGGTGTGCCCCATCTGTGAGCTGAGC 60
Db 1096 ATGTGAGCAAGGCGAGAGAGCTTTCAACGGGGTGTGCCCCATCTGTGAGCTGAGC 1155
QY 61 GGCACGTAAGCGGCAAAAGTTGAGCGGTTCGCGGAGCCTGAGGCGATGCCACTTAC 120
Db 1156 GGCACGTAAGCGGCAAAAGTTGAGCGGTTCGCGGAGCCTGAGGCGATGCCACTTAC 1215
QY 121 GGCAGCTGACCTCTGAAGTTTATCTGACCAACCGGCAAGCTGCCCCCTGCGCCACC 180
Db 1216 GGCAGCTGACCTCTGAAGTTTATCTGACCAACCGGCAAGCTGCCCCCTGCGCCACC 1275
QY 181 CTCTGACCACTTCACTGACCTGAGGCGTGTGAGCTTCAAGCGCTACCCCGACCATATGAG 240
Db 1276 CTCTGACCACTTCACTGACCTGAGGCGTGTGAGCTTCAAGCGCTACCCCGACCATATGAG 1335
QY 241 CAGACGACCTTCTTCAAGTCCGCGCATGCGCGAAGCTTCGTCAGAGCGGACCATCTTC 300
Db 1336 CAGACGACCTTCTTCAAGTCCGCGCATGCGCGAAGCTTCGTCAGAGCGGACCATCTTC 1395
QY 301 TTCAAGGACGAGCGCAACTCAAGACCGCGCGAGGTGAAGTTGAGGCGGACACCTTG 360
Db 1396 TTCAAGGACGAGCGCAACTCAAGACCGCGCGAGGTGAAGTTGAGGCGGACACCTTG 1455
QY 361 GTGAACCGCATGAGCTGAGAGGCGCATGACTTCAAGAGAGGACGCAACATCTGCGGCGAC 420
Db 1456 GTGAACCGCATGAGCTGAGAGGCGCATGACTTCAAGAGAGGACGCAACATCTGCGGCGAC 1515
QY 421 AAGCTGAGTAACAATCAACAGCAACAAGCTCTATTCATGAGCGGCAAGCAAGAAC 480
Db 1516 AAGCTGAGTAACAATCAACAGCAACAAGCTCTATTCATGAGCGGCAAGCAAGAAC 1575
QY 481 GGCATCAAGTGAATCTCAAGATCCGCCAACAATCGAGAGCGCAACGCTGAGCTGCC 540
Db 1576 GGCATCAAGTGAATCTCAAGATCCGCCAACAATCGAGAGCGCAACGCTGAGCTGCC 1635
QY 541 GACCACTTACAGAGCAACACCCCATGCGGAGCGGCGCGCTGCTGCGCCGACACAC 600
Db 1636 GACCACTTACAGAGCAACACCCCATGCGGAGCGGCGCGCTGCTGCGCCGACACAC 1695
QY 601 TACCTGAGCAAGCGGCGGCGCTGAGCAAGAACCCCAAGAGAGGCGGATCATATGTC 660
Db 1696 TACCTGAGCAAGCGGCGGCGCTGAGCAAGAACCCCAAGAGAGGCGGATCATATGTC 1755
QY 661 CTGCTGAGGTTCTGACCGCGCGCGGAGTCACTCTGCGCATGAGCGAGCTGTACATTA 719
Db 1756 CTGCTGAGGTTCTGACCGCGCGCGGAGTCACTCTGCGCATGAGCGAGCTGTACATTA 1814

RESULT 15
US-09-417-197-64
; Sequence 64, Application US/09417197
; Patent No. 6518021
; GENERAL INFORMATION:
; APPLICANT: Ole Thastrup, et al.
; TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An I
; TITLE OF INVENTION: On A Cellular Response
; FILE REFERENCE: 3759-0110P
; CURRENT APPLICATION NUMBER: US/09/417,197
; CURRENT FILING DATE: 1999-10-07
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 64
; LENGTH: 1821
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: p38-BGFP fusion
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1818)
US-09-417-197-64

Query Match 84.4%; Score 717.4; DB 4; Length 1821;
Best Local Similarity 99.9%; Pred. No. 3e-134; Indels 0; Gaps 0;
Matches 718; Conservative 0; Mismatches 1;

QY 1 ATGTTAGCAAGGCGAGAGCTTTCAACGGGGTGTGCCCCATCTGTGAGCTGAGC 60
Db 1102 ATGTGAGCAAGGCGAGAGAGCTTTCAACGGGGTGTGCCCCATCTGTGAGCTGAGC 1161
QY 61 GGCACGTAAGCGGCAAAAGTTGAGCGGTTCGCGGAGCCTGAGGCGATGCCACTTAC 120
Db 1162 GGCACGTAAGCGGCAAAAGTTGAGCGGTTCGCGGAGCCTGAGGCGATGCCACTTAC 1221
QY 121 GGCAGCTGACCTCTGAAGTTTATCTGACCAACCGGCAAGCTGCCCCCTGCGCCACC 180
Db 1222 GGCAGCTGACCTCTGAAGTTTATCTGACCAACCGGCAAGCTGCCCCCTGCGCCACC 1281
QY 181 CTCTGACCACTTCACTGACCTGAGGCGTGTGAGCTTCAAGCGCTACCCCGACCATATGAG 240
Db 1282 CTCTGACCACTTCACTGACCTGAGGCGTGTGAGCTTCAAGCGCTACCCCGACCATATGAG 1341
QY 241 CAGACGACCTTCTTCAAGTCCGCGCATGCGCGAAGCTTCGTCAGAGCGGACCATCTTC 300
Db 1342 CAGACGACCTTCTTCAAGTCCGCGCATGCGCGAAGCTTCGTCAGAGCGGACCATCTTC 1401
QY 301 TTCAAGGACGAGCGCAACTCAAGACCGCGCGAGGTGAAGTTGAGGCGGACACCTTG 360
Db 1402 TTCAAGGACGAGCGCAACTCAAGACCGCGCGAGGTGAAGTTGAGGCGGACACCTTG 1461
QY 361 GTGAACCGCATGAGCTGAGAGGCGCATGACTTCAAGAGAGGACGCAACATCTGCGGCGAC 420
Db 1462 GTGAACCGCATGAGCTGAGAGGCGCATGACTTCAAGAGAGGACGCAACATCTGCGGCGAC 1521
QY 421 AAGCTGAGTAACAATCAACAGCAACAAGCTCTATTCATGAGCGGCAAGCAAGAAC 480
Db 1522 AAGCTGAGTAACAATCAACAGCAACAAGCTCTATTCATGAGCGGCAAGCAAGAAC 1581
QY 481 GGCATCAAGTGAATCTCAAGATCCGCCAACAATCGAGAGCGGCAACGCTGAGCTGCC 540
Db 1582 GGCATCAAGTGAATCTCAAGATCCGCCAACAATCGAGAGCGGCAACGCTGAGCTGCC 1641
QY 541 GACCACTTACAGAGCAACACCCCATGCGGAGCGGCGCGCTGCTGCGCCGACACAC 600
Db 1642 GACCACTTACAGAGCAACACCCCATGCGGAGCGGCGCGCTGCTGCGCCGACACAC 1701
QY 601 TACCTGAGCAAGCGGCGGCGCTGAGCAAGAACCCCAAGAGAGGCGGATCATATGTC 660
Db 1702 TACCTGAGCAAGCGGCGGCGCTGAGCAAGAACCCCAAGAGAGGCGGATCATATGTC 1761

Mon May 17 09:20:38 2004

us-09-931-232-2.rn1

Page 10

QY	661	CTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGAA	719
Db	1762	CTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTA	1820

Search completed: May 15, 2004, 05:22:04
Job time : 90 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2004, 04:33:32 ; Search time 465 Seconds

(without alignments)
8295.358 Million cell updates/sec

Title: US-09-931-232-2

Perfect score: 850
1 atggtgagcaaggcgagga.....tagatcaatgtgtagatgc 850

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2947324 seqs, 2269024515 residues

Total number of hits satisfying chosen parameters: 5894648

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	850	100.0	850	13	US-09-931-232-2
2	850	100.0	4862	15	US-10-161-403-87
3	721.2	84.8	10417	15	US-10-152-040-28
4	720.6	84.8	6418	16	US-10-408-456-2
5	720	84.7	12789	13	US-10-666-778-9
6	719.4	84.6	8528	15	US-10-239-804-11
7	719.2	84.6	5713	9	US-09-963-206B-5
8	719.2	84.6	5713	9	US-09-966-976A-5
9	719	84.6	8531	12	US-10-429-608A-2
10	719	84.6	8531	15	US-10-134-643-5
11	719	84.6	8531	16	US-10-408-456-1
12	719	84.6	8531	17	US-10-716-725-2
13	718.6	84.5	1190	11	US-09-927-876-98
14	718.6	84.5	1190	12	US-10-457-047-98

15	718.6	84.5	1190	16	US-10-360-149-98	Sequence 98, Appl
16	718.6	84.5	1377	15	US-10-204-724-1	Sequence 1, Appl
17	718.6	84.5	4151	13	US-10-421-285-15	Sequence 15, Appl
18	718.6	84.5	4733	10	US-09-797-496B-1	Sequence 1, Appl
19	718.6	84.5	4733	15	US-10-177-590-1	Sequence 1, Appl
20	718.6	84.5	4944	14	US-10-001-189-55	Sequence 52, Appl
21	718.6	84.5	4952	14	US-10-001-189-52	Sequence 50, Appl
22	718.6	84.5	9013	14	US-10-001-189-50	Sequence 3, Appl
23	718.6	84.5	9941	15	US-10-243-553-3	Sequence 2, Appl
24	718.6	84.5	9941	15	US-10-243-816-2	Sequence 2, Appl
25	718.6	84.5	9941	15	US-10-243-820-9	Sequence 9, Appl
26	718.6	84.5	9941	15	US-10-243-817-9	Sequence 9, Appl
27	718.6	84.5	9941	15	US-10-319-341-3	Sequence 3, Appl
28	718.2	84.5	5162	10	US-09-815-981-13	Sequence 13, Appl
29	718.2	84.5	5162	10	US-09-815-979-13	Sequence 13, Appl
30	718.2	84.5	5162	13	US-10-428-653-13	Sequence 13, Appl
31	718.2	84.5	5162	15	US-10-235-119-13	Sequence 13, Appl
32	718.2	84.5	5162	15	US-10-161-403-26	Sequence 26, Appl
33	718.2	84.5	5162	15	US-10-086-745-13	Sequence 13, Appl
34	718	84.5	5100	10	US-09-954-483A-14	Sequence 14, Appl
35	718	84.5	5510	15	US-10-161-403-71	Sequence 71, Appl
36	718	84.5	6100	15	US-10-314-861-36	Sequence 36, Appl
37	718	84.5	6115	15	US-10-314-861-34	Sequence 34, Appl
38	718	84.5	6119	15	US-10-161-403-126	Sequence 126, Appl
39	718	84.5	6121	15	US-10-314-861-38	Sequence 38, Appl
40	718	84.5	6133	15	US-10-314-861-32	Sequence 32, Appl
41	718	84.5	6151	15	US-10-314-861-30	Sequence 30, Appl
42	718	84.5	6169	15	US-10-314-861-28	Sequence 28, Appl
43	718	84.5	6248	15	US-10-314-861-15	Sequence 15, Appl
44	718	84.5	7600	15	US-10-161-403-115	Sequence 115, Appl
45	718	84.5	7631	15	US-10-161-403-116	Sequence 116, Appl

ALIGNMENTS

RESULT 1
US-09-931-232-2
; Sequence 2, Application US/09931232
; Publication No. US20020058274A1
GENERAL INFORMATION:
APPLICANT: Li, Xiangqiang
APPLICANT: Kahn, Steve
TITLE OF INVENTION: Rapidly Degrading GFP-Fusion Proteins
TITLE OF INVENTION: and Methods of Use
FILE REFERENCE: CLON075CON
CURRENT APPLICATION NUMBER: US/09/931,232
CURRENT FILING DATE: 2001-08-15
PRIOR APPLICATION NUMBER: 09/364,946
PRIOR FILING DATE: 1999-07-30
PRIOR APPLICATION NUMBER: 09/191,233
PRIOR FILING DATE: 1998-11-13
PRIOR APPLICATION NUMBER: 09/062,102
PRIOR FILING DATE: 1998-04-17
PRIOR APPLICATION NUMBER: 60/060,855
PRIOR FILING DATE: 1997-10-02
NUMBER OF SEQ ID NOS: 3
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 850
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Sequence of the EGFP-MOD422_461 fusion
US-09-931-232-2

Query Match 100.0%; Score 850; DB 13; Length 850;
Best Local Similarity 100.0%; Pred. No. 6; 4e-190;
Matches 850; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 ATGTTGAGCAAGGCGAGGAGCTGTTCACCGGGGTGTGCGCCATCCGTGTGACTGTGAC 60
|||||


```

Db      1  ATGGTGAAGCAAGGCGAGAGAGCTGTTCAACCGGGGTGGTGGCCATCTGTGTGAGCTGAGC 60
QY      61  GGGGAGCTAAACGGCCCAAGATTCAAGGTGTCCGGGAGGGCGAGGGCGCATGCACTTAC 120
Db      61  GGGGAGCTAAACGGCCCAAGATTCAAGGTGTCCGGGAGGGCGAGGGCGCATGCACTTAC 120
QY      121  GGGCAAGTGAACCTTGAAGTTTCACTGTGACCAACCGGCAAGCTGCCCTGCTGCGCCACC 180
Db      121  GGGCAAGTGAACCTTGAAGTTTCACTGTGACCAACCGGCAAGCTGCCCTGCTGCGCCACC 180
QY      181  CTGTGTGAACCACTTGAACCTTGAACCTTGAACCTTGAACCTTGAACCTTGAACCTTGA 240
Db      181  CTGTGTGAACCACTTGAACCTTGAACCTTGAACCTTGAACCTTGAACCTTGAACCTTGA 240
QY      241  CAGCAGCACTTCTTCAAGTCCGCGCATGCGCGAGGAGGCTTCAAGAGGCGCAACATCTTC 300
Db      241  CAGCAGCACTTCTTCAAGTCCGCGCATGCGCGAGGAGGCTTCAAGAGGCGCAACATCTTC 300
QY      301  TTCAAGAGCAGCAGGCAACTTCAAGAACCCGCGCGAGGAGGAGGAGGAGGAGGAGGAGG 360
Db      301  TTCAAGAGCAGCAGGCAACTTCAAGAACCCGCGCGAGGAGGAGGAGGAGGAGGAGGAGG 360
QY      361  GTGAACCGCATTCGAGCTGAAGGGGATCGACTTCAAGAGAGAGGAGGAGGAGGAGGAGGAG 420
Db      361  GTGAACCGCATTCGAGCTGAAGGGGATCGACTTCAAGAGAGAGGAGGAGGAGGAGGAGGAG 420
QY      421  AAGCTGAGTGAACCTTCAAGAACCCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
Db      421  AAGCTGAGTGAACCTTCAAGAACCCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
QY      481  GGCATCAAGGTGAACCTTCAAGAACCCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540
Db      481  GGCATCAAGGTGAACCTTCAAGAACCCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540
QY      541  GACCACTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
Db      541  GACCACTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
QY      601  TACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
Db      601  TACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
QY      661  CTGCTGAGAGTTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
Db      661  CTGCTGAGAGTTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
QY      721  CTTAGCATAGGCTTCCCGCGGAGGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
Db      721  CTTAGCATAGGCTTCCCGCGGAGGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
QY      781  TGTGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Db      781  TGTGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
QY      841  GTGTGAATGC 850
Db      841  GTGTGAATGC 850

```

RESULT 2

```

US-10-161-403-87
; Sequence 87, Application US/10161403
; Publication No. US20030119104A1
; GENERAL INFORMATION:
; APPLICANT: Perez, Edward
; APPLICANT: Perkins, Carl
; APPLICANT: Lindenbaum, Michael
; APPLICANT: Greene, Amy
; APPLICANT: Leung, Josephine
; APPLICANT: Fleming, Elena
; APPLICANT: Stewart, Sandra
; APPLICANT: Shellard, Joan
; TITLE OF INVENTION: CHROMOSOME-BASED PLATFORMS

```

```

; FILE REFERENCE: 24601-420
; CURRENT APPLICATION NUMBER: US/10161,403
; CURRENT FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: 60/294,758
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: 60/366,891
; PRIOR FILING DATE: 2002-03-21
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 87
; LENGTH: 4862
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: pDzeGFP-IN plasmid from Clontech
US-10-161-403-87

```

```

Query Match      100.0%; Score 850; DB 15; Length 4862;
Best Local Similarity 100.0%; Pred. No. 7.5e-150;
Matches 850; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1  ATGGTGAAGCAAGGCGAGAGAGCTGTTCAACCGGGGTGGTGGCCATCTGTGTGAGCTGAGC 60
Db      619  ATGGTGAAGCAAGGCGAGAGAGCTGTTCAACCGGGGTGGTGGCCATCTGTGTGAGCTGAGC 738
QY      61  GGGCAGTGAACGGCCCAAGTTCAAGCTGTCCGGCGAGGGCGAGGCGATGCCACTTAC 120
Db      739  GGGCAGTGAACGGCCCAAGTTCAAGCTGTCCGGCGAGGGCGAGGCGATGCCACTTAC 798
QY      121  GGGCAAGTGAACCTTGAAGTTTCACTGTGACCAACCGGCAAGCTGCCCTGCTGCGCCACC 180
Db      799  GGGCAAGTGAACCTTGAAGTTTCACTGTGACCAACCGGCAAGCTGCCCTGCTGCGCCACC 858
QY      181  CTGTGTGAACCACTTGAACCTTGAACCTTGAACCTTGAACCTTGAACCTTGAACCTTGA 240
Db      859  CTGTGTGAACCACTTGAACCTTGAACCTTGAACCTTGAACCTTGAACCTTGAACCTTGA 918
QY      241  CAGCAGCACTTCTTCAAGTCCGCGCATGCGCGAGGAGGCTTCAAGAGGCGCAACATCTTC 300
Db      919  CAGCAGCACTTCTTCAAGTCCGCGCATGCGCGAGGAGGCTTCAAGAGGCGCAACATCTTC 978
QY      301  TTCAAGAGCAGCAGGCAACTTCAAGAACCCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
Db      979  TTCAAGAGCAGCAGGCAACTTCAAGAACCCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAG 1038
QY      361  GTGAACCGCATTCGAGCTGAAGGGGATCGACTTCAAGAGAGAGAGAGAGAGAGAGAGAGAG 420
Db      1039  GTGAACCGCATTCGAGCTGAAGGGGATCGACTTCAAGAGAGAGAGAGAGAGAGAGAGAGAG 1098
QY      421  AAGCTGAGTGAACCTTCAAGAACCCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
Db      1099  AAGCTGAGTGAACCTTCAAGAACCCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1158
QY      481  GGCATCAAGGTGAACCTTCAAGAACCCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540
Db      1159  GGCATCAAGGTGAACCTTCAAGAACCCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1218
QY      541  GACCACTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
Db      1219  GACCACTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1278
QY      601  TACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
Db      1279  TACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1338
QY      661  CTGCTGAGAGTTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
Db      1339  CTGCTGAGAGTTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1398
QY      721  CTTAGCATAGGCTTCCCGCGGAGGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
Db      1399  CTTAGCATAGGCTTCCCGCGGAGGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1458

```

QY	1	ATGTGACCAAGAGGCGAGAGCTGTTCAACCGAGGAGTGAGCCATCTGGTGTGAGCTGAC	60
Dp	743	ATGTGACCAAGGCGAGAGAGCTGTTCAACCGAGGAGTGAGCCATCTGGTGTGAGCTGAC	802
QY	61	GCGCAGCTAAACGSCCAAGTTCAAGCGTGCAGGCGAGAGGCGAGCGATCCACTAC	120
Dp	803	GCGCAGCTAAACGSCCAAGTTCAAGCGTGTCCGGCGAGAGGCGAGAGCGATCCACTAC	862
QY	121	GGCAAGCTGACCCCTGAAGTTCACTTGACCAACCGGCGAAGCTGCCCTGTGCTTGGCCAC	180
Dp	863	GGCAAGCTGACCCCTGAAGTTCACTTGACCAACCGGCGAAGCTGCCCTGTGCTTGGCCAC	922
QY	181	CTCGTGACCAACCTGACCTAGCGGCGTGAGTGCTTCAACCGCTACCCGACCAACATGAAG	240
Dp	923	CTCGTGACCAACCTGACCTAGCGGCGTGAGTGCTTCAACCGCTACCCGACCAACATGAAG	982
QY	241	CAGACGCACTTCTTCAAGTCCGCCATGCCCCGAAGGCTACGTCAGAGAGCGACCATCTTC	300
Dp	983	CAGACGCACTTCTTCAAGTCCGCCATGCCCCGAAGGCTACGTCAGAGAGCGACCATCTTC	1042
QY	301	TTCAAGGACGACGGCACTACAGAACCCGCGCGAGGGAAGTTGAGAGGCGACACCTGTG	360
Dp	1043	TTCAAGGACGACGGCACTACAGAACCCGCGCGAGGGAAGTTGAGAGGCGACACCTGTG	1102
QY	361	GTGAACCGCATCGAGCTGAAGGACATGCACTTCAAGAGGAGACGGCAACATCTGTGGGAC	420
Dp	1103	GTGAACCGCATCGAGCTGAAGGACATGCACTTCAAGAGGAGACGGCAACATCTGTGGGAC	1162
QY	421	AAGGTGAGATCAACTACAGCCACAAGCTTTATCAAGGCGGACCAAGCATGAAGAAC	480
Dp	1163	AAGGTGAGATCAACTACAGCCACAAGCTTTATCAAGGCGGACCAAGCATGAAGAAC	1222

QY	1	ATGTTGGCAAGGGCGAGAGCTGTTACCGGGGTGTGTGCCATCTGATGTGAGCTGCAC	60
Db	2690	ATGTTGAGCAAGGGCGAGAGCTGTTACCGGGGTGTGTGCCATCTGATGTGAGCTGCAC	2744
QY	61	GGGACGCTAAACGGCCCAAGTTGAGCTGTCCGGCGAGGGCGAGGGCGATGCCACTTAC	120
Db	2750	GGGACGCTAAACGGCCCAAGTTGAGCTGTCCGGCGAGGGCGAGGGCGATGCCACTTAC	2805
QY	121	GGCAAGCTGAACCTGGAAGTTCACTGCAACCAACGGCAAGCTGCCCTGTGGCCAC	180
Db	2810	GGCAAGCTGAACCTGGAAGTTCACTGCAACCAACGGCAAGCTGCCCTGTGGCCAC	2865
QY	181	CTGTGACCAACCTGACCTTACGGGTGACAGTGTTCACCGCGCTACCCCGACCAATGAG	240
Db	2870	CTGTGACCAACCTGACCTTACGGGTGACAGTGTTCACCGCGCTACCCCGACCAATGAG	2929
QY	241	CAGCACAGCTTCTTCAAGTCGCGCATGCCGAAAGCTACGTCAGAGAGCGCACCATCTTC	300
Db	2930	CAGCACAGCTTCTTCAAGTCGCGCATGCCGAAAGCTACGTCAGAGAGCGCACCATCTTC	2985
QY	301	TTTCAAGAGCAGCGCAACTACAAGACCGGCGCGAGGTGAAATTCGAGGGGAGACACCTTG	360
Db	2990	TTTCAAGAGCAGCGCAACTACAAGACCGGCGCGAGGTGAAATTCGAGGGGAGACACCTTG	3049
QY	361	GTGAACCGCATCGAGCTGAAGGGCGATCGACTTCAAGAGAGAAGCGCAACATCTTGGGGCAC	420

Db 3050 GTGAACCGATCGAGCTGAAGGGGCGATCGCTTCAAGAGAGCGCAACATCTCGGGGCGAC 3109
QY 421 AAGCTGAGTACAACTACAAAGCCAAAGCTTATATATGAGCGAGAGAGAAAG 480
Db 3110 AAGCTGAGTACAACTACAAAGCCAAAGCTTATATATGAGCGAGAGAGAAAG 3169
QY 481 GGCATCAAGGTGAATCTTCAAGATCGGCAACATCGAGAGCGGAGCGTGCAGTGC 540
Db 3170 GGCATCAAGGTGAATCTTCAAGATCGGCAACATCGAGAGCGGAGCGTGCAGTGC 3229
QY 541 GACCACTACAGAGAAACCCCGATCGGAGAGCGGCGGCTGCTGCTGCCCAACAC 600
Db 3230 GACCACTACAGAGAAACCCCGATCGGAGAGCGGCGGCTGCTGCTGCCCAACAC 3289
QY 601 TACCTGAGACCCAGTCCCGCTGAGCAAAAGACCCCAAGAGAGCGGATCATGATG 660
Db 3290 TACCTGAGACCCAGTCCCGCTGAGCAAAAGACCCCAAGAGAGCGGATCATGATG 3349
QY 661 CTGCTGAGGTTGCTGACCGCGCGGAGTCACTCTCGGATGAGCGAGCTGTACAAAG 720
Db 3350 CTGCTGAGGTTGCTGACCGCGCGGAGTCACTCTCGGATGAGCGAGCTGTACAAAG 3409
QY 721 CTGAGCCATGGCTCCCGCGGAGGAGGAGAGAGAT 759
Db 3410 AGCGCGCGGAGCTGTAGAGTCACTGCAAGAAATTCGAT 3448

RESULT 5
US-10-666-778-9/C
Sequence 9, Application US/10666778
Publication No. US20040064849A1
GENERAL INFORMATION:
APPLICANT: Goossens, Alain
APPLICANT: Inze, Dirk
TITLE OF INVENTION: THE USE OF GENES ENCODING MEMBERANE TRANSPORTER PUMPS TO STIMULATE
FILE REFERENCE: DJ/ABC/V082
CURRENT APPLICATION NUMBER: US/10/666,778
CURRENT FILING DATE: 2003-09-18
PRIOR APPLICATION NUMBER: EP01201407.2
PRIOR FILING DATE: 2001-04-18
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn version 3.1
SEQ ID NO 9
LENGTH: 12789
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: vector pK7W2D
NAME/KEY: misc_feature
LOCATION: (4772)..(4772)
OTHER INFORMATION: n can be any base
US-10-666-778-9

Query Match 84.7%; Score 720; DB 13; Length 12789;
Best Local Similarity 100.0%; Pred. No. 2.5e-159;
Matches 720; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGGTGAGCAAGGGGCGAGGAGCTTACCGGGGTGGTCCCATCTGTGTGACAGCTGAC 60
Db 5797 ATGGTGAAGAGGGGCGAGGAGCTTACCGGGGTGGTCCCATCTGTGTGACAGCTGAC 5738
QY 61 GGGGAGTAAAGCGGCACAAAGTTCAAGCTGTCCGAGGAGGCGAGGCGATGCCACTTAC 120
Db 5737 GGGGAGTAAAGCGGCACAAAGTTCAAGCTGTCCGAGGAGGCGAGGCGATGCCACTTAC 5678
QY 121 GGGAGACTGACCTGTAAGTTCACTGTGACCAACCGGAGAGGCTGCGCTGAGCCACC 180
Db 5677 GGGAGACTGACCTGTAAGTTCACTGTGACCAACCGGAGAGGCTGCGCTGAGCCACC 5618
QY 181 CTGTGACCAACCTGACCTTACGAGCGGTGCTTACGCGGCTAACCCGACCAATGAAG 240

Db 5617 CTGTGACCAACCTGACCTTACGAGCGGTGAGTGTCTTCAACCCCTACCCGACCAATGAAG 5558
QY 241 CAGCAGACTTCTTAAATGTCGCCATGACCCGAAAGGCTGCTCGAGAGCGGACCATCTTC 300
Db 5557 CAGCAGACTTCTTAAATGTCGCCATGACCCGAAAGGCTGCTCGAGAGCGGACCATCTTC 5498
QY 301 TTCAAGAGCAGACGGCACTTACAAAGACCCGCGCCGAGGTGAATTCGAGGCGACACCC 360
Db 5497 TTCAAGAGCAGACGGCACTTACAAAGACCCGCGCCGAGGTGAATTCGAGGCGACACCC 5438
QY 361 GTGAACCGATTCGAGCTGAAGGGCATTCGACTTCAAGAGAGCGGCAACATCTTGGGCGAC 420
Db 5437 GTGAACCGATTCGAGCTGAAGGGCATTCGACTTCAAGAGAGCGGCAACATCTTGGGCGAC 5378
QY 421 AAGCTGAGTACAACTTACAAAGAGCAACGCTTATATATGAGCGAGAGAGAAAG 480
Db 5377 AAGCTGAGTACAACTTACAAAGAGCAACGCTTATATATGAGCGAGAGAGAAAG 5318
QY 481 GGCATCAAGGTGAATCTTCAAGATCGGCAACATCGAGAGCGGAGCGTGCAGCTGC 540
Db 5317 GGCATCAAGGTGAATCTTCAAGATCGGCAACATCGAGAGCGGAGCGTGCAGCTGC 5258
QY 541 GACCACTACAGCAAAACCCCATCGGAGAGCGGCGGCTGCTGCTGCCCAACAC 600
Db 5257 GACCACTACAGCAAAACCCCATCGGAGAGCGGCGGCTGCTGCTGCCCAACAC 5198
QY 601 TACCTGAGACCCAGTCCCGCTGAGCAAAAGACCCCAAGAGAGCGGATCATGATG 660
Db 5197 TACCTGAGACCCAGTCCCGCTGAGCAAAAGACCCCAAGAGAGCGGATCATGATG 5138
QY 661 CTGCTGAGGTTGCTGACCGCGCGGAGTCACTCTCGGATGAGCGAGCTGTACAAAG 720
Db 5137 CTGCTGAGGTTGCTGACCGCGCGGAGTCACTCTCGGATGAGCGAGCTGTACAAAG 5078

RESULT 6
US-10-239-804-11
Sequence 11, Application US/10239804
Publication No. US20030053991A1
GENERAL INFORMATION:
APPLICANT: Oxford Biomedica (UK) Limited
APPLICANT: Kingsman, Alan J
APPLICANT: Maden, Malcolm
APPLICANT: Corcoran, Jonathan PT
TITLE OF INVENTION: Factor
FILE REFERENCE: P009156MOCTH
CURRENT APPLICATION NUMBER: US/10/239,804
CURRENT FILING DATE: 2002-09-23
PRIOR APPLICATION NUMBER: PCT/GB00/01211
PRIOR FILING DATE: 2000-03-30
PRIOR APPLICATION NUMBER: GB 0024300.6
PRIOR FILING DATE: 2000-10-04
NUMBER OF SEQ ID NOS: 73
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 11
LENGTH: 8528
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: PONY8G 5' cPPT
OTHER INFORMATION: POS delCTS E1AV vector genome plasmid
US-10-239-804-11

Query Match 84.6%; Score 719.4; DB 15; Length 8528;
Best Local Similarity 97.2%; Pred. No. 3.4e-159;
Matches 732; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 1 AAGTGAAGCAAGGGGCGAGAGCTTACCGGGGTGGTCCCATCTGTGTGAGCTGAC 60
Db 2868 AAGTGAAGCAAGGGGCGAGAGCTTACCGGGGTGGTCCCATCTGTGTGAGCTGAC 2927
QY 61 GGGGAGTAAAGCGGCACAAAGTTCAAGCTGTCCGAGGAGGCGAGGAGATGCCACTTAC 120

Db 2928 GCGCGCTGAAACGGCCAAAGTTTCAGCGTGTCCGCGCGAGGCGAGGCGGATGCCACTTAC 2987
Qy 121 GCGAAGCTGACCCCTGAAGTTTATCTGACACACCGGCAAGCTGCCCCGTCGCCGCCACC 180
Db 2988 GCGAAGCTGACCCCTGAAGTTTATCTGACACACCGGCAAGCTGCCCCGTCGCCGCCACC 3047
Qy 181 CTCGTGACCAACCCCTGACCTTACGCGCGTGAAGTCTTCAAGCCGTCACCCGACCACTATGAAG 240
Db 3048 CTCGTGACCAACCCCTGACCTTACGCGCGTGAAGTCTTCAAGCCGTCACCCGACCACTATGAAG 3107
Qy 241 CAGCAGCACTTCTTCAAGTCCGCAATGCCCGGAGGCTTACGTCAGAGGCGACCATCTTC 300
Db 3108 CAGCAGCACTTCTTCAAGTCCGCAATGCCCGGAGGCTTACGTCAGAGGCGACCATCTTC 3167
Qy 301 TTCAAGGAGCGAGCGCACTTCAAGACCGCGCGGAGGTTGAAGTGAAGGCGACCCCTG 360
Db 3168 TTCAAGGAGCGAGCGCACTTCAAGACCGCGCGGAGGTTGAAGTGAAGGCGACCCCTG 3227
Qy 361 GTGAACCGCATCGAGCTGAAGGCGATCGACTTCAAGAGAGAGCGGCAATCTCTGCGGCGAC 420
Db 3228 GTGAACCGCATCGAGCTGAAGGCGATCGACTTCAAGAGAGAGCGGCAATCTCTGCGGCGAC 3287
Qy 421 AAGCTGAGTCACTTCAAGACCGCAACGCTTATATCATGCGGACCAAGCAAGAAC 480
Db 3288 AAGCTGAGTCACTTCAAGACCGCAACGCTTATATCATGCGGACCAAGCAAGAAC 3347
Qy 481 GGCATCAAGTGAATCTTCAAGATCCGCAACATCGAGAGAGCGGAGGTGAGTCCGC 540
Db 3348 GGCATCAAGTGAATCTTCAAGATCCGCAACATCGAGAGAGCGGAGGTGAGTCCGC 3407
Qy 541 GACCACTACCAAGCAACACCCCATCGCGAGCGGCGCCGCTGCTGCGCCGACCAACAC 600
Db 3408 GACCACTACCAAGCAACACCCCATCGCGAGCGGCGCCGCTGCTGCGCCGACCAACAC 3467
Qy 601 TACCTGAGACCCCACTCCGCTGAGCAAAAGCCCAAGAGAGCGGATCACTGCTC 660
Db 3468 TACCTGAGACCCCACTCCGCTGAGCAAAAGCCCAAGAGAGCGGATCACTGCTC 3527
Qy 661 CTGCTGAGTTCGTGACCGCGCGGAGTCACTCTCGGATGAGAGAGTGTGACAAAG 720
Db 3528 CTGCTGAGTTCGTGACCGCGCGGAGTCACTCTCGGATGAGAGAGTGTGACAAAG 3587
Qy 721 CTTAGCCATGCTTCCCGCGGAGGTGAGAGAG 753
Db 3588 AGCGGCGCGACTTGAAGTGAAGTCTGAGAGG 3620

RESULT 7
US-09-963-206B-5
; Sequence 5, Application US/09963206B
; Patent No. US20020123076A1
; GENERAL INFORMATION:
; APPLICANT: Ferrick, David A.
; APPLICANT: Swift, Susan B.
; APPLICANT: Armstrong, Randall
; APPLICANT: Fox, Bryan
; TITLE OF INVENTION: Methods and Compositions for Screening for Modulators and Ige Syn
; FILE REFERENCE: A-66038-3/RMS/JTD/DIR
; CURRENT APPLICATION NUMBER: US/09/963,206B
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 09/076,624
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 5
; LENGTH: 5713
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-963-206B-5

Query Match 84.6%; Score 719.2; DB 9; Length 5713;
Best Local Similarity 98.9%; Pred. No. 3,6e-159;
Matches 724; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
Qy 1 ATGTGAGCAAGGGCGAGAGCTTTCACCGGGGTGTGTCATCTGTGTGAGCTGAGC 60
Db 2150 ATGTGAGCAAGGGCGAGAGCTTTCACCGGGGTGTGTCATCTGTGTGAGCTGAGC 2209
Qy 61 GCGCAAGTAAACCGGCAAGATTGAGCGTGTGCGGCAAGGCGAGGCGATGCCACTTAC 120
Db 2210 GCGCAAGTAAACCGGCAAGATTGAGCGTGTGCGGCAAGGCGAGGCGATGCCACTTAC 2269
Qy 121 GCGAAGCTGACCCCTGAAGTTTATCTGACACACCGGCAAGCTGCCCCGTCGCCGCCACC 180
Db 2270 GCGAAGCTGACCCCTGAAGTTTATCTGACACACCGGCAAGCTGCCCCGTCGCCGCCACC 2329
Qy 181 CTCGTGACCAACCCCTGACCTTACGCGCGTGAAGTCTTCAAGCCGTCACCCGACCACTATGAAG 240
Db 2330 CTCGTGACCAACCCCTGACCTTACGCGCGTGAAGTCTTCAAGCCGTCACCCGACCACTATGAAG 2389
Qy 241 CAGCAGCACTTCTTCAAGTCCGCAATGCCCGGAGGCTTACGTCAGAGGCGACCATCTTC 300
Db 2390 CAGCAGCACTTCTTCAAGTCCGCAATGCCCGGAGGCTTACGTCAGAGGCGACCATCTTC 2449
Qy 301 TTCAAGGAGCGAGCGCACTTCAAGACCGCGCGGAGGTTGAAGTGAAGGCGACCACTTC 360
Db 2450 TTCAAGGAGCGAGCGCACTTCAAGACCGCGCGGAGGTTGAAGTGAAGGCGACCACTTC 2509
Qy 361 GTGAACCGCATCGAGCTGAAGGCGATCGACTTCAAGAGAGAGCGGCAATCTCTGCGGCGAC 420
Db 2510 GTGAACCGCATCGAGCTGAAGGCGATCGACTTCAAGAGAGAGCGGCAATCTCTGCGGCGAC 2569
Qy 421 AAGCTGAGTCACTTCAAGACCGCAACGCTTATATCATGCGGACCAAGCAAGAAC 480
Db 2570 AAGCTGAGTCACTTCAAGACCGCAACGCTTATATCATGCGGACCAAGCAAGAAC 2629
Qy 481 GGCATCAAGTGAATCTTCAAGATCCGCAACATCGAGAGAGCGGATCACTGCTC 540
Db 2630 GGCATCAAGTGAATCTTCAAGATCCGCAACATCGAGAGAGCGGATCACTGCTC 2689
Qy 541 GACCACTACCAAGCAACACCCCATCGCGAGCGGCGCCGCTGCTGCGCCGACCAACAC 600
Db 2690 GACCACTACCAAGCAACACCCCATCGCGAGCGGCGCCGCTGCTGCGCCGACCAACAC 2749
Qy 601 TACCTGAGACCCCACTCCGCTGAGCAAAAGCCCAAGAGAGCGGATCACTGCTC 660
Db 2750 TACCTGAGACCCCACTCCGCTGAGCAAAAGCCCAAGAGAGCGGATCACTGCTC 2809
Qy 661 CTGCTGAGTTCGTGACCGCGCGGAGTCACTCTCGGATGAGAGAGTGTGACAAAG 720
Db 2810 CTGCTGAGTTCGTGACCGCGCGGAGTCACTCTCGGATGAGAGAGTGTGACAAAG 2869
Qy 721 CTTAGCCATGCTC 732
Db 2870 TTGAGAGTGC 2881

RESULT 8
US-09-966-976A-5
; Sequence 5, Application US/09966976A
; Patent No. US2002016849A1
; GENERAL INFORMATION:
; APPLICANT: Ferrick, David A.
; APPLICANT: Swift, Susan B.
; APPLICANT: Armstrong, Randall
; APPLICANT: Fox, Bryan
; TITLE OF INVENTION: Methods and Compositions for Screening for Modulators and Ige Syn
; FILE REFERENCE: A-66038-4/RMS/JTD/DIR
; CURRENT APPLICATION NUMBER: US/09/966,976A
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 09/076,624

PRIOR FILING DATE: 1998-05-12
NUMBER OF SEQ ID NOS: 19
SOFTWARE: PatentIn version 3.1
SEQ ID NO 5
LENGTH: 5713
TYPE: DNA
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: synthetic
US-09-966-976A-5

Query Match 84.6%; Score 719.2; DB 9; Length 5713;
Best Local Similarity 98.3%; Pred. No. 3.0e-159;
Matches 724; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 ATGGTGAAGCAAGGCGGAGAGCTGTTACCGGAGTGTGCTCCATCTGTGAGCTGAGAC 60
DB 2150 ATGGTGAAGCAAGGCGGAGAGCTGTTACCGGAGTGTGCTCCATCTGTGAGCTGAGAC 2209
QY 61 GGGGAGCTAAAGGCGGCAAGTTGAGCTGTCCGCGAGGCGAGGCGGAGTCCACTTAC 120
DB 2210 GGGGAGCTAAAGGCGGCAAGTTGAGCTGTCCGCGAGGCGAGGCGGAGTCCACTTAC 2269
QY 121 GGGGAGCTGAGCTGAGTTCATCTGACCAACCGGAGAGTGTCCGCGAGGCGGAGTCCACTTAC 180
DB 2270 GGGGAGCTGAGCTGAGTTCATCTGACCAACCGGAGAGTGTCCGCGAGGCGGAGTCCACTTAC 2329
QY 181 CTGCTGAGCAACCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 240
DB 2330 CTGCTGAGCAACCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 2389
QY 241 CAGCAGACTTCTTCAAGTCCGCAATGCCGAGGCTGAGCTGAGCTGAGCTGAGCTGAG 300
DB 2390 CAGCAGACTTCTTCAAGTCCGCAATGCCGAGGCTGAGCTGAGCTGAGCTGAGCTGAG 2449
QY 301 TTCAAGGAGCAAGGCGGCAAGTTGAGCTGTCCGCGAGGCGAGGCGGAGTCCACTTAC 360
DB 2450 TTCAAGGAGCAAGGCGGCAAGTTGAGCTGTCCGCGAGGCGAGGCGGAGTCCACTTAC 2509
QY 361 GTGAACCGGATCGAGCTGAGGAGGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 420
DB 2510 GTGAACCGGATCGAGCTGAGGAGGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 2569
QY 421 AAGCTGAGTGAACCTTCAAGTCCGCAATGCCGAGGCTGAGCTGAGCTGAGCTGAG 480
DB 2570 AAGCTGAGTGAACCTTCAAGTCCGCAATGCCGAGGCTGAGCTGAGCTGAGCTGAG 2629
QY 481 GGCATCAAGGTGAACCTTCAAGTCCGCAATGCCGAGGCTGAGCTGAGCTGAGCTGAG 540
DB 2630 GGCATCAAGGTGAACCTTCAAGTCCGCAATGCCGAGGCTGAGCTGAGCTGAGCTGAG 2689
QY 541 GACCACTACCAAGGAGCAACCTTCAAGTCCGCAATGCCGAGGCTGAGCTGAGCTGAG 600
DB 2690 GACCACTACCAAGGAGCAACCTTCAAGTCCGCAATGCCGAGGCTGAGCTGAGCTGAG 2749
QY 601 TACCTGAGCAACCTTCAAGTCCGCAATGCCGAGGCTGAGCTGAGCTGAGCTGAG 660
DB 2750 TACCTGAGCAACCTTCAAGTCCGCAATGCCGAGGCTGAGCTGAGCTGAGCTGAG 2809
QY 661 CTGCTGAGTGAACCTTCAAGTCCGCAATGCCGAGGCTGAGCTGAGCTGAGCTGAG 720
DB 2810 CTGCTGAGTGAACCTTCAAGTCCGCAATGCCGAGGCTGAGCTGAGCTGAGCTGAG 2869
QY 721 CTGCTGAGTGAACCTTCAAGTCCGCAATGCCGAGGCTGAGCTGAGCTGAG 780
DB 2870 CTGCTGAGTGAACCTTCAAGTCCGCAATGCCGAGGCTGAGCTGAGCTGAG 2881

RESULT 9
US-10-429-608A-2
Sequence 2, Application US/10429608A
Publication No. US20040071675A1
GENERAL INFORMATION:

APPLICANT: MAZARAKIS, NICHOLAS
APPLICANT: AZOUZ, MIMOUN
TITLE OF INVENTION: VECTOR SYSTEM
FILE REFERENCE: 674523-2017
CURRENT APPLICATION NUMBER: US/10/429, 608A
PRIORITY FILING DATE: 2003-05-05
PRIORITY APPLICATION NUMBER: PCT/GB01/04866
PRIORITY FILING DATE: 2001-11-02
PRIORITY APPLICATION NUMBER: GB 0122238.9
PRIORITY FILING DATE: 2001-09-14
PRIORITY APPLICATION NUMBER: GB 0102339.9
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: GB 0026943.1
PRIORITY FILING DATE: 2000-11-03
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 8531
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
OTHER INFORMATION: nucleotide construct PONTAG sequence
US-10-429-608A-2

Query Match 84.6%; Score 719; DB 12; Length 8531;
Best Local Similarity 97.3%; Pred. No. 4.2e-159;
Matches 731; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 1 ATGGTGAAGCAAGGCGGAGAGCTGTTACCGGAGTGTGCTCCATCTGTGAGCTGAGAC 60
DB 2690 ATGGTGAAGCAAGGCGGAGAGCTGTTACCGGAGTGTGCTCCATCTGTGAGCTGAGAC 2749
QY 61 GGGGAGCTAAAGGCGGCAAGTTGAGCTGTCCGCGAGGCGAGGCGGAGTCCACTTAC 120
DB 2750 GGGGAGCTAAAGGCGGCAAGTTGAGCTGTCCGCGAGGCGAGGCGGAGTCCACTTAC 2809
QY 121 GGGGAGCTGAGCTGAGTTCATCTGACCAACCGGAGAGTGTCCGCGAGGCGGAGTCCACTTAC 180
DB 2810 GGGGAGCTGAGCTGAGTTCATCTGACCAACCGGAGAGTGTCCGCGAGGCGGAGTCCACTTAC 2869
QY 181 CTGCTGAGCAACCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 240
DB 2870 CTGCTGAGCAACCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 2929
QY 241 CAGCAGACTTCTTCAAGTCCGCAATGCCGAGGCTGAGCTGAGCTGAGCTGAGCTGAG 300
DB 2930 CAGCAGACTTCTTCAAGTCCGCAATGCCGAGGCTGAGCTGAGCTGAGCTGAGCTGAG 2989
QY 301 TTCAAGGAGCAAGGCGGCAAGTTGAGCTGTCCGCGAGGCGAGGCGGAGTCCACTTAC 360
DB 3070 TTCAAGGAGCAAGGCGGCAAGTTGAGCTGTCCGCGAGGCGAGGCGGAGTCCACTTAC 3049
QY 361 GTGAACCGGATCGAGCTGAGGAGGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 420
DB 3050 GTGAACCGGATCGAGCTGAGGAGGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 3109
QY 421 AAGCTGAGTGAACCTTCAAGTCCGCAATGCCGAGGCTGAGCTGAGCTGAGCTGAG 480
DB 3110 AAGCTGAGTGAACCTTCAAGTCCGCAATGCCGAGGCTGAGCTGAGCTGAGCTGAG 3169
QY 481 GGCATCAAGGTGAACCTTCAAGTCCGCAATGCCGAGGCTGAGCTGAGCTGAGCTGAG 540
DB 3170 GGCATCAAGGTGAACCTTCAAGTCCGCAATGCCGAGGCTGAGCTGAGCTGAGCTGAG 3229
QY 541 GACCACTACCAAGGAGCAACCTTCAAGTCCGCAATGCCGAGGCTGAGCTGAGCTGAG 600
DB 3230 GACCACTACCAAGGAGCAACCTTCAAGTCCGCAATGCCGAGGCTGAGCTGAGCTGAG 3289
QY 601 TACCTGAGCAACCTTCAAGTCCGCAATGCCGAGGCTGAGCTGAGCTGAGCTGAG 660
DB 3290 TACCTGAGCAACCTTCAAGTCCGCAATGCCGAGGCTGAGCTGAGCTGAGCTGAG 3349

QY 661 CTGCTGAGATTCTGTAACCGCCGCGGATCATCTTCGGCATGAGACGAGCTGTAAAG 720
DB 3350 CTGCTGAGATTCTGTAACCGCCGCGGATCATCTTCGGCATGAGACGAGCTGTAAAG 3409
QY 721 CTGAGCATGCTTCCCGCGGAGGTGAG 751
DB 3410 AGCGCCGCGACTCTTAGAGTGCAGCTGCAG 3440

RESULT 10

US-10-134-643-5
; Sequence 5, Application US/10134643
; Publication No. US20030113898A1
; GENERAL INFORMATION:
; APPLICANT: OLSEN, JOHN C.
; APPLICANT: MITROPHANOUS, KYRIACOS ANDREOU
; APPLICANT: ROHLI, JONATHAN
; APPLICANT: KINGSMAN, ALAN JOHN
; APPLICANT: ELLIARD, FIONA MARGARET
; TITLE OF INVENTION: METHODS FOR PRODUCING HIGH TITRE VECTORS AND
; FILE REFERENCE: 078883-0148
; CURRENT APPLICATION NUMBER: US/10/134,643
; PRIOR FILING DATE: 2002-08-30
; PRIOR APPLICATION NUMBER: 60/287,048
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 5
; LENGTH: 8531
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: nucleic acid pONY8g, vector genome plasmid
US-10-134-643-5

Query Match 84.6%; Score 719; DB 15; Length 8531;
Best Local Similarity 97.3%; Pred. No. 4.2e-159;
Matches 731; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 1 ATGTGAGCAAGGGGAGAGAGCTGTTACCGGGGTGTGCCATCTGTCAGCTGAGC 60
DB 2690 ATGTGAGCAAGGGGAGAGAGCTGTTACCGGGGTGTGCCATCTGTCAGCTGAGC 2749
QY 61 GGCGACGTAAGCGGCACAAAGTTCAAGCGTGTCCGGCGAGGGGCGATGCCACTAC 120
DB 2750 GGCGACGTAAGCGGCACAAAGTTCAAGCGTGTCCGGCGAGGGGCGATGCCACTAC 2809
QY 121 GGCAAGCTGACCTTGAAGTTCACTGCAACCGGCAAGCTGCGCTGCTGCGCCAC 180
DB 2810 GGCAAGCTGACCTTGAAGTTCACTGCAACCGGCAAGCTGCGCTGCTGCGCCAC 2869
QY 181 CTGCGACCACTTCAAGTCCGCACTGAGAGCTTCAAGCTTCAAGCTTCAAGCTTCA 240
DB 2870 CTGCGACCACTTCAAGTCCGCACTGAGAGCTTCAAGCTTCAAGCTTCAAGCTTCA 2929
QY 241 CAGCAAGCTTCAAGTCCGCACTGAGAGCTTCAAGCTTCAAGCTTCAAGCTTCA 300
DB 2930 CAGCAAGCTTCAAGTCCGCACTGAGAGCTTCAAGCTTCAAGCTTCAAGCTTCA 2989
QY 301 TTCAAGGACGACGCGCACTAAGACCGGCGCGAGGTGAAGTTGAGGGCGACACCTG 360
DB 2990 TTCAAGGACGACGCGCACTAAGACCGGCGCGAGGTGAAGTTGAGGGCGACACCTG 3049
QY 361 GTGAACCGCATCGAGCTGAAGGCGATCGACTTCAAGGAGGACGCGAATCTCGGGGAC 420
DB 3050 GTGAACCGCATCGAGCTGAAGGCGATCGACTTCAAGGAGGAGCGCAATCTCGGGGAC 3109
QY 421 AAGCTGAGTAACTAACAACAAGCCAAAGCTTATCATGCGCGACAGAGAAAGAC 480
DB 3110 AAGCTGAGTAACTAACAACAAGCCAAAGCTTATCATGCGCGACAGAGAAAGAC 3169

QY 481 GGCATCAAGGTGAATTCAAGATCCGCCAACAATCGAGAGGAGCGGTGAGCTCGCC 540
DB 3170 GGCATCAAGGTGAATTCAAGATCCGCCAACAATCGAGAGGAGCGGTGAGCTCGCC 3229
QY 541 GACCATCAAGCAAGAACACCCCAATCGCGAGCGGCCCGCTGTCGCCCAACAACAC 600
DB 3230 GACCATCAAGCAAGAACACCCCAATCGCGAGCGGCCCGCTGTCGCCCAACAACAC 3289
QY 601 TACCTGAGCAACCACTCGCCCTGAGCAAGACCCCAAGAGAGCGGATCAATGTC 660
DB 3290 TACCTGAGCAACCACTCGCCCTGAGCAAGACCCCAAGAGAGCGGATCAATGTC 3349
QY 661 CTGCTGAGATTGTAAACCGCGCGGATCACTTCGAGATGAGAGCTGTAAAGAG 720
DB 3350 CTGCTGAGATTGTAAACCGCGCGGATCACTTCGAGATGAGAGCTGTAAAGAG 3409
QY 721 CTGAGCATGCTTCCCGCGGAGGTGAG 751
DB 3410 AGCGCCGCGACTCTTAGAGTGCAGCTGCAG 3440

RESULT 11

US-10-408-456-1
; Sequence 1, Application US/10408456
; Publication No. US20040013648A1
; GENERAL INFORMATION:
; APPLICANT: Oxford Biomedica (UK) Limited
; APPLICANT: KINGSMAN, et al., Alan John
; TITLE OF INVENTION: Vector System
; FILE REFERENCE: 674523-2016
; CURRENT APPLICATION NUMBER: US/10/408,456
; PRIOR FILING DATE: 2003-04-08
; PRIOR APPLICATION NUMBER: PCT/GB01/04433
; PRIOR FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: 0024550.6
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: Patent In version 3.0
; SEQ ID NO 1
; LENGTH: 8531
; TYPE: DNA
; ORGANISM: Equine Infectious Anemia Virus
US-10-408-456-1

Query Match 84.6%; Score 719; DB 15; Length 8531;
Best Local Similarity 97.3%; Pred. No. 4.2e-159;
Matches 731; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 1 ATGTGAGCAAGGGGAGAGAGCTGTTACCGGGGTGTGCCATCTGTCAGCTGAGC 60
DB 2690 ATGTGAGCAAGGGGAGAGAGCTGTTACCGGGGTGTGCCATCTGTCAGCTGAGC 2749
QY 61 GGCGACGTAAGCGGCACAAAGTTCAAGCGTGTCCGGCGAGGGGCGATGCCACTAC 120
DB 2750 GGCGACGTAAGCGGCACAAAGTTCAAGCGTGTCCGGCGAGGGGCGATGCCACTAC 2809
QY 121 GGCAAGCTGACCTTGAAGTTCACTGCAACCGGCAAGCTGCGCTGCTGCGCCAC 180
DB 2810 GGCAAGCTGACCTTGAAGTTCACTGCAACCGGCAAGCTGCGCTGCTGCGCCAC 2869
QY 181 CTGCGACCACTTCAAGTCCGCACTGAGAGCTTCAAGCTTCAAGCTTCAAGCTTCA 240
DB 2870 CTGCGACCACTTCAAGTCCGCACTGAGAGCTTCAAGCTTCAAGCTTCAAGCTTCA 2929
QY 241 CAGCAAGCTTCAAGTCCGCACTGAGAGCTTCAAGCTTCAAGCTTCAAGCTTCA 300
DB 2930 CAGCAAGCTTCAAGTCCGCACTGAGAGCTTCAAGCTTCAAGCTTCAAGCTTCA 2989
QY 301 TTCAAGGACGACGCGCACTAAGACCGGCGCGAGGTGAAGTTGAGGGCGACACCTG 360
DB 2990 TTCAAGGACGACGCGCACTAAGACCGGCGCGAGGTGAAGTTGAGGGCGACACCTG 3049
QY 361 GTGAACCGCATCGAGCTGAAGGCGATCGACTTCAAGGAGGACGCGAATCTCGGGGAC 420

Db 3050 GTAAACCGCATGAGCTGAAAGGGCATCGACTTCAAGAGAGACGGCAACATCTCGGGCAGC 3109
QY 421 AAGCTGAGTACAACTACACACAGCCCAACAGTCTATATCATGGCCGCAAGAGAAAGAC 480
Db 3110 AAGCTGAGTACAACTACACACAGCCCAACAGTCTATATCATGGCCGCAAGAGAAAGAC 3169
QY 481 GGCATCAAGGTGAACCTTCAAGATCCGCCACACATCGAGAGACGGACGCTGACCTGACC 540
Db 3170 GGCATCAAGGTGAACCTTCAAGATCCGCCACACATCGAGAGACGGACGCTGACCTGACC 3229
QY 541 GACCACTACCAAGAGAAACACCCCATTCGGGAGACGGCCCGCTGCTGCTGCCGCAACAC 600
Db 3230 GACCACTACCAAGAGAAACACCCCATTCGGGAGACGGCCCGCTGCTGCTGCCGCAACAC 3289
QY 601 TACCTGAGACACCAAGTCCGCCCTGAGCAAGACCCCAACAGAGAGCGCATCATGCTC 660
Db 3290 TACCTGAGACACCAAGTCCGCCCTGAGCAAGACCCCAACAGAGAGCGCATCATGCTC 3349
QY 661 CTGCTGAGAGTTCGTGACCCGCCCGCGGATCATCTCGGCATGAGACGAGCTGTACAAAG 720
Db 3350 CTGCTGAGAGTTCGTGACCCGCCCGCGGATCATCTCGGCATGAGACGAGCTGTACAAAG 3409
QY 721 CTTAAGCCATGCTTCCCGCCGAGGTGAGG 751
Db 3410 AGCGGCCGCGACTCTAGAGTTCGACCTGCAAG 3440

RESULT 12
US-10-716-725-2
; Sequence 2, Application US/10716725
; Publication No. US2004007613A1
; GENERAL INFORMATION:
; APPLICANT: MAZARAKIS, NICHOLAS
; APPLICANT: AZZOUZ, MINOUN
; APPLICANT: KINGSMAN, SUSAN
; TITLE OF INVENTION: VECTOR SYSTEM
; FILE REFERENCE: 674523-2017.1
; CURRENT APPLICATION NUMBER: US/10/716,725
; PRIOR FILING DATE: 2003-11-19
; PRIOR APPLICATION NUMBER: 10/429,608
; PRIOR FILING DATE: 2003-05-05
; PRIOR APPLICATION NUMBER: PCT/GB03/00426
; PRIOR FILING DATE: 2003-10-03
; PRIOR APPLICATION NUMBER: GB 0223076.1
; PRIOR FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: GB 0228314.1
; PRIOR FILING DATE: 2002-12-04
; PRIOR APPLICATION NUMBER: GB 0318213.6
; PRIOR FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: PCT/GB01/04866
; PRIOR FILING DATE: 2001-11-02
; PRIOR APPLICATION NUMBER: GB 0122238.9
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: GB 0102339.9
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: GB 0026943.1
; PRIOR FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 8531
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: nucleotide construct ponyrbg sequence
US-10-716-725-2

Query Match 84.6%; Score 719; DB 17; Length 8531;
Best Local Similarity 97.3%; Pred. No. 4.2e-159;
Matches 731; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 1 ATGTGAGCAAGGCGAGAGCTGTTACCGGGATGATGCCCATCTGCTCGAGCTGAC 60
Db 2690 ATGTGAGCAAGGCGAGAGCTGTTACCGGGATGATGCCCATCTGCTCGAGCTGAC 2749
QY 61 GGCAGCTTAAAGGCGCAAAAGTTCAAGCTGCTCCGCGAGAGGCGATGCCATCTAC 120
Db 2750 GGCAGCTTAAAGGCGCAAAAGTTCAAGCTGCTCCGCGAGAGGCGATGCCATCTAC 2809
QY 121 GGCAGCTTAAAGGCGCAAAAGTTCAAGCTGCTCCGCGAGAGGCGATGCCATCTAC 180
Db 2810 GGCAGCTTAAAGGCGCAAAAGTTCAAGCTGCTCCGCGAGAGGCGATGCCATCTAC 2869
QY 181 CTGTCGACCAACCTGACCTTACCGCGCTGACAGTCTTCAAGCGCTAACCCGACCAATGAA 240
Db 2870 CTGTCGACCAACCTGACCTTACCGCGCTGACAGTCTTCAAGCGCTAACCCGACCAATGAA 2929
QY 241 CAGCAGACCTTTCTTAAGTCCGCCATGCCCGAAGGCTACGTCAGAGAGCGCATCTTTC 300
Db 2930 CAGCAGACCTTTCTTAAGTCCGCCATGCCCGAAGGCTACGTCAGAGAGCGCATCTTTC 2989
QY 301 TTCAAGAGACAGGCGCAACTTACAAAGACCGCGCGCGAGGTGAAGTTGAGAGGCGACACCTG 360
Db 2990 TTCAAGAGACAGGCGCAACTTACAAAGACCGCGCGCGAGGTGAAGTTGAGAGGCGACACCTG 3049
QY 361 GTGAACCGCATGAGCTGAGGCGCATGCACTTCAAGAGAGACGCAACATCTGAGGCGAC 420
Db 3050 GTGAACCGCATGAGCTGAGGCGCATGCACTTCAAGAGAGACGCAACATCTGAGGCGAC 3109
QY 421 AAGCTGAGTACAACTACACACAGCCCAACAGTCTATATCATGGCCGCAAGAGAAAGAC 480
Db 3110 AAGCTGAGTACAACTACACACAGCCCAACAGTCTATATCATGGCCGCAAGAGAAAGAC 3169
QY 481 GGCATCAAGGTGAACCTTCAAGATCCGCCACACATCGAGAGACGGACGCTGACCTGACC 540
Db 3170 GGCATCAAGGTGAACCTTCAAGATCCGCCACACATCGAGAGACGGACGCTGACCTGACC 3229
QY 541 GACCACTACCAAGAGAAACACCCCATTCGGGAGACGGCCCGCTGCTGCTGCCGCAACAC 600
Db 3230 GACCACTACCAAGAGAAACACCCCATTCGGGAGACGGCCCGCTGCTGCTGCCGCAACAC 3289
QY 601 TACCTGAGACACCAAGTCCGCCCTGAGCAAGACCCCAACAGAGAGCGCATCATGCTC 660
Db 3290 TACCTGAGACACCAAGTCCGCCCTGAGCAAGACCCCAACAGAGAGCGCATCATGCTC 3349
QY 661 CTGCTGAGAGTTCGTGACCCGCCCGCGGATCATCTCGGCATGAGACGAGCTGTACAAAG 720
Db 3350 CTGCTGAGAGTTCGTGACCCGCCCGCGGATCATCTCGGCATGAGACGAGCTGTACAAAG 3409
QY 721 CTTAAGCCATGCTTCCCGCCGAGGTGAGG 751
Db 3410 AGCGGCCGCGACTCTAGAGTTCGACCTGCAAG 3440

RESULT 13
US-09-927-876-98
; Sequence 98, Application US/09927876
; Publication No. US2004005554A1
; GENERAL INFORMATION:
; APPLICANT: El Tayar, Nabil
; APPLICANT: Campbell, Robert K
; APPLICANT: Kelton, Christie A
; APPLICANT: He, Chaomei
; TITLE OF INVENTION: No. US2004005554A1 Glycoproteins and Methods of Use Thereof
; FILE REFERENCE: 20993-003
; CURRENT APPLICATION NUMBER: US/09/927,876
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/225,035
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: 60/202,724
; PRIOR FILING DATE: 2000-05-08
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 98

```

: GENERAL INFORMATION:
: APPLICANT: El Tayar, Nabil
: APPLICANT: Campbell, Robert K
: APPLICANT: Kelton, Christie A
: APPLICANT: He, Chaomei

```

Db	1051	TACCTGAGCAACCAGTCCG
QY	661	CTGCTGAGTTCGTGACCG
Db	1111	CTGCTGAGTTCGTGACCG

QY 721 CTTAGCCATGGCT 733
DB 1171 AGCGCCCGGACT 1183

RESULT 15

US-10-360-149-98
Sequence 98, Application US/10360149
Publication No. US20030219786A1
GENERAL INFORMATION:
APPLICANT: El Tavar, Nabli
APPLICANT: Campbell, Robert K
APPLICANT: Kelton, Christie A
APPLICANT: He, Chaomei
TITLE OF INVENTION: No. US20030219786A1 Glycoproteins and Methods of Use Thereof
FILE REFERENCE: 20993-003
CURRENT APPLICATION NUMBER: US/10/360,149
CURRENT FILING DATE: 2003-02-06
PRIOR APPLICATION NUMBER: US/09/927,876
PRIOR FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: 60/225,035
PRIOR FILING DATE: 2000-08-11
PRIOR APPLICATION NUMBER: 60/202,724
NUMBER OF SEQ ID NOS: 107
SOFTWARE: Patentn Ver. 2.1
SEQ ID NO 98
LENGTH: 1190
TYPE: DNA
ORGANISM: Artificial Sequence
OTHER INFORMATION: Description of Artificial Sequence: Fusion Protein
US-10-360-149-98

Query Match 84.5%; Score 718.6; DB 16; Length 1190;
Best Local Similarity 98.8%; Pred. No. 4.4e-159;
Matches 724; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 ATGTGAGCAAGGCGGAGAGCTGTTACCGGGGTGTGCCCCATCTGTGAGCTGAGC 60
DB 451 ATGTGAGCAAGGCGGAGAGCTGTTACCGGGGTGTGCCCCATCTGTGAGCTGAGC 510
QY 61 GCGGAGCTTAAAGGCGGAGAGCTGTTACCGGGGTGTGCCCCATCTGTGAGCTGAGC 120
DB 511 GCGGAGCTTAAAGGCGGAGAGCTGTTACCGGGGTGTGCCCCATCTGTGAGCTGAGC 570
QY 121 GCGGAGCTTAAAGGCGGAGAGCTGTTACCGGGGTGTGCCCCATCTGTGAGCTGAGC 180
DB 571 GCGGAGCTTAAAGGCGGAGAGCTGTTACCGGGGTGTGCCCCATCTGTGAGCTGAGC 630
QY 181 CTGTGAGCAAGGCGGAGAGCTGTTACCGGGGTGTGCCCCATCTGTGAGCTGAGC 240
DB 631 CTGTGAGCAAGGCGGAGAGCTGTTACCGGGGTGTGCCCCATCTGTGAGCTGAGC 690
QY 241 CAGCAGCACTTCTTAAGTCCGCGGAGAGCTGTTACCGGGGTGTGCCCCATCTGTGAGC 300
DB 691 CAGCAGCACTTCTTAAGTCCGCGGAGAGCTGTTACCGGGGTGTGCCCCATCTGTGAGC 750
QY 301 TTCAAGGAGAGGAGAGCTGTTACCGGGGTGTGCCCCATCTGTGAGCTGAGC 360
DB 751 TTCAAGGAGAGGAGAGCTGTTACCGGGGTGTGCCCCATCTGTGAGCTGAGC 810
QY 361 GTGAACCGGATGAGCTGTTACCGGGGTGTGCCCCATCTGTGAGCTGAGC 420
DB 811 GTGAACCGGATGAGCTGTTACCGGGGTGTGCCCCATCTGTGAGCTGAGC 870
QY 421 AAGCTGAGTAACTTAACTGAGAGAGAGCTGTTACCGGGGTGTGCCCCATCTGTGAGC 480
DB 871 AAGCTGAGTAACTTAACTGAGAGAGAGCTGTTACCGGGGTGTGCCCCATCTGTGAGC 930
QY 481 GGCATCAAGGTAACTTAACTGAGAGAGAGCTGTTACCGGGGTGTGCCCCATCTGTGAGC 540

DB 931 GGCATCAAGGTAACTTAACTGAGAGAGAGCTGTTACCGGGGTGTGCCCCATCTGTGAGC 990
QY 541 GACCACTAACGAGAGAGAGCTGTTACCGGGGTGTGCCCCATCTGTGAGCTGAGC 600
DB 991 GACCACTAACGAGAGAGAGCTGTTACCGGGGTGTGCCCCATCTGTGAGCTGAGC 1050
QY 601 TACCTGAGAGAGAGAGCTGTTACCGGGGTGTGCCCCATCTGTGAGCTGAGC 660
DB 1051 TACCTGAGAGAGAGAGCTGTTACCGGGGTGTGCCCCATCTGTGAGCTGAGC 1110
QY 661 CTGCTGAGAGAGAGAGCTGTTACCGGGGTGTGCCCCATCTGTGAGCTGAGC 720
DB 1111 CTGCTGAGAGAGAGAGCTGTTACCGGGGTGTGCCCCATCTGTGAGCTGAGC 1170
QY 721 CTTAGCCATGGCT 733
DB 1171 AGCGCCCGGACT 1183

Search completed: May 15, 2004, 06:32:39
Job time : 470 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2004, 03:23:12 ; Search time 3573 Seconds
(without alignments)
8370.613 Million cell updates/sec

Title: US-09-931-232-2

Perfect score: 850
Sequence: 1 atgtgtgagcaagcgagga.....tagatcatgtgtatgc 850

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 37577330 seqs, 17593059518 residues

Total number of hits satisfying chosen parameters: 75154660

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents NA Main:*

- 1: /cgn2_6/ptodata/2/pna/US0933A_COMB.seq:*
- 2: /cgn2_6/ptodata/2/pna/US0933A_COMB.seq:*
- 3: /cgn2_6/ptodata/2/pna/US0933A_COMB.seq:*
- 4: /cgn2_6/ptodata/2/pna/US0933A_COMB.seq:*
- 5: /cgn2_6/ptodata/2/pna/US0933A_COMB.seq:*
- 6: /cgn2_6/ptodata/2/pna/US0933A_COMB.seq:*
- 7: /cgn2_6/ptodata/2/pna/US0933A_COMB.seq:*
- 8: /cgn2_6/ptodata/2/pna/US0933A_COMB.seq:*
- 9: /cgn2_6/ptodata/2/pna/US0933A_COMB.seq:*
- 10: /cgn2_6/ptodata/2/pna/US0933A_COMB.seq:*
- 11: /cgn2_6/ptodata/2/pna/US0933A_COMB.seq:*
- 12: /cgn2_6/ptodata/2/pna/US0933A_COMB.seq:*
- 13: /cgn2_6/ptodata/2/pna/US0933A_COMB.seq:*
- 14: /cgn2_6/ptodata/2/pna/US0933A_COMB.seq:*
- 15: /cgn2_6/ptodata/2/pna/US0933A_COMB.seq:*
- 16: /cgn2_6/ptodata/2/pna/US0933A_COMB.seq:*
- 17: /cgn2_6/ptodata/2/pna/US0933A_COMB.seq:*
- 18: /cgn2_6/ptodata/2/pna/US0933A_COMB.seq:*
- 19: /cgn2_6/ptodata/2/pna/US0933A_COMB.seq:*
- 20: /cgn2_6/ptodata/2/pna/US0933A_COMB.seq:*
- 21: /cgn2_6/ptodata/2/pna/US0933A_COMB.seq:*
- 22: /cgn2_6/ptodata/2/pna/US0933A_COMB.seq:*
- 23: /cgn2_6/ptodata/2/pna/US0933A_COMB.seq:*
- 24: /cgn2_6/ptodata/2/pna/US0933A_COMB.seq:*
- 25: /cgn2_6/ptodata/2/pna/US0933A_COMB.seq:*
- 26: /cgn2_6/ptodata/2/pna/US0933A_COMB.seq:*
- 27: /cgn2_6/ptodata/2/pna/US0933A_COMB.seq:*
- 28: /cgn2_6/ptodata/2/pna/US0933A_COMB.seq:*
- 29: /cgn2_6/ptodata/2/pna/US0933A_COMB.seq:*
- 30: /cgn2_6/ptodata/2/pna/US0933A_COMB.seq:*
- 31: /cgn2_6/ptodata/2/pna/US0933A_COMB.seq:*
- 32: /cgn2_6/ptodata/2/pna/US0933A_COMB.seq:*
- 33: /cgn2_6/ptodata/2/pna/US0933A_COMB.seq:*
- 34: /cgn2_6/ptodata/2/pna/US0933A_COMB.seq:*
- 35: /cgn2_6/ptodata/2/pna/US0933A_COMB.seq:*
- 36: /cgn2_6/ptodata/2/pna/US0933A_COMB.seq:*
- 37: /cgn2_6/ptodata/2/pna/US0933A_COMB.seq:*
- 38: /cgn2_6/ptodata/2/pna/US0933A_COMB.seq:*
- 39: /cgn2_6/ptodata/2/pna/US0933A_COMB.seq:*
- 40: /cgn2_6/ptodata/2/pna/US0933A_COMB.seq:*
- 41: /cgn2_6/ptodata/2/pna/US0933A_COMB.seq:*
- 42: /cgn2_6/ptodata/2/pna/US0933A_COMB.seq:*
- 43: /cgn2_6/ptodata/2/pna/US0933A_COMB.seq:*

- 44: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 45: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 46: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 47: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 48: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 49: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 50: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 51: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 52: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 53: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 54: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 55: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 56: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 57: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 58: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 59: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 60: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 61: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 62: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 63: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 64: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 65: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 66: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 67: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 68: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 69: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 70: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 71: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 72: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 73: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 74: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 75: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 76: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 77: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 78: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 79: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 80: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 81: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 82: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 83: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 84: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 85: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 86: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 87: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 88: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 89: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 90: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 91: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 92: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 93: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 94: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 95: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 96: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 97: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 98: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 99: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 100: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 101: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 102: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 103: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 104: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 105: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 106: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 107: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 108: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 109: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
- 110: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysts of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	850	100.0	850	1	PCT-US98-24333-2	Sequence 2, Appli
2	850	100.0	850	15	US-09-191-233-2	Sequence 2, Appli
3	850	100.0	850	15	US-09-191-233-2	Sequence 2, Appli
4	850	100.0	850	19	US-09-136-089-2	Sequence 2, Appli
5	850	100.0	850	38	US-09-931-222-2	Sequence 2, Appli
6	850	100.0	4862	1	PCT-US102-17452-87	Sequence 87, Appli
7	850	100.0	4862	46	US-10-161-403-87	Sequence 87, Appli
8	846	99.5	4163	1	PCT-US99-28461-12	Sequence 12, Appli
9	846	99.5	4163	16	US-09-206-887-12	Sequence 12, Appli
10	846	99.5	4300	1	PCT-US99-28451-13	Sequence 13, Appli
11	846	99.5	4300	16	US-09-206-887-13	Sequence 13, Appli
12	845	99.4	885	49	US-10-333-733-21	Sequence 21, Appli
13	845	99.4	6748	1	PCT-US03-28111-6	Sequence 3, Appli
14	845	99.4	7969	1	PCT-US03-28111-6	Sequence 3, Appli
15	845	99.4	10417	46	US-10-152-040-28	Sequence 28, Appli
16	848	84.8	6418	51	US-10-408-455-2	Sequence 2, Appli
17	720.6	84.7	6418	51	US-10-408-456-2	Sequence 2, Appli
18	720	84.7	12789	53	US-10-666-778-9	Sequence 9, Appli
19	719.4	84.6	8525	49	US-10-334-235-17	Sequence 17, Appli
20	719.4	84.6	8528	47	US-10-239-804-11	Sequence 11, Appli
21	719.2	84.6	5713	14	US-09-076-624-5	Sequence 5, Appli
22	719.2	84.6	5713	14	US-09-076-624-5	Sequence 5, Appli
23	719.2	84.6	5713	30	US-09-712-821-5	Sequence 5, Appli
24	719.2	84.6	5713	41	US-09-962-206B-5	Sequence 5, Appli
25	719.2	84.6	5713	41	US-09-962-247A-5	Sequence 5, Appli
26	719.2	84.6	5713	41	US-09-966-976A-5	Sequence 5, Appli
27	719	84.6	8531	45	US-10-138-643-5	Sequence 5, Appli
28	719	84.6	8531	51	US-10-408-455-1	Sequence 1, Appli
29	719	84.6	8531	51	US-10-408-456-1	Sequence 1, Appli
30	719	84.6	8531	51	US-10-422-608A-2	Sequence 2, Appli
31	719	84.6	8531	54	US-10-726-725-2	Sequence 2, Appli
32	719	84.6	10292	44	US-10-088-076-47	Sequence 47, Appli
33	719	84.6	10292	44	US-10-088-076-48	Sequence 48, Appli
34	719	84.6	10384	44	US-10-088-076-45	Sequence 45, Appli
35	719	84.6	10384	44	US-10-088-076-46	Sequence 46, Appli
36	718.6	84.5	761	35	US-09-869-588-25	Sequence 25, Appli
37	718.6	84.5	761	54	US-10-731-988-35	Sequence 35, Appli
38	718.6	84.5	1190	38	US-09-927-876-58	Sequence 98, Appli
39	718.6	84.5	1190	50	US-10-360-149-58	Sequence 98, Appli
40	718.6	84.5	1190	52	US-10-457-047-98	Sequence 98, Appli
41	718.6	84.5	1377	47	US-10-204-724-1	Sequence 1, Appli
42	718.6	84.5	4151	51	US-10-421-285-15	Sequence 15, Appli
43	718.6	84.5	4412	53	US-10-677-777-3	Sequence 3, Appli
44	718.6	84.5	4733	15	US-09-177-300-9	Sequence 1, Appli
45	718.6	84.5	4733	24	US-09-575-309-9	Sequence 9, Appli

ALIGNMENTS

```

RESULT 1
PCT-US98-24323-2
; Sequence 2 Application PC/TUS9824323A
; GENERAL INFORMATION:
; APPLICANT: Kaln, Steve
; APPLICANT: Li, Xiangdang
; TITLE OF INVENTION: Rapidly Degrading GFP-Fusion Proteins and Methods
; TITLE OF INVENTION: of use
; FILE REFERENCE: D6100PCT
; CURRENT APPLICATION NUMBER: PCT/US98/24323A
; CURRENT FILING DATE: 1998-11-13
; EARLIER APPLICATION NUMBER: US 09/062,102
; EARLIER FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 3
; SEQ ID NO 2
; LENGTH: 850
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:

```

```

; OTHER INFORMATION: DNA sequence which encodes the EGFP-MODC422-461
; OTHER INFORMATION: fusion protein.
PCT-US98-24323-2

```

Query Match	100.0%;	Score 850;	DB 1;	Length 850;
Best Local Similarity	100.0%;	Pred. No. 2.5e-134;		
Matches 850; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

QY	1	ATGGTAGCAAGGGCGAAGAGCTGTTCAACCGAGGAGTGAGCCATATCTGGTAGACTGGAC	60
Db	1	ATGGTAGCAAGGGCGAAGAGCTGTTCAACCGAGGAGTGAGCCATATCTGGTAGACTGGAC	60
QY	61	GGCGACGTAAACGGCCACAAGTTTCAGCGTGTCGCGCGAAGGGCGAGGGCGATGCCATTAC	120
Db	61	GGCGACGTAAACGGCGACAAGTTTCAGCGTGTCGCGCGAAGGGCGAGGGCGATGCCATTAC	120
QY	121	GGCAACGTAAACCGCTAAAGTTCACTCGACCAACCGGCAAGCTGCGCCGTGAGCCCTGCGCCAC	180
Db	121	GGCAACGTAAACCGCTAAAGTTCACTCGACCAACCGGCAAGCTGCGCCGTGAGCCCTGCGCCAC	180
QY	181	CTTCGTGACCAACCGCTGACCTTACCGGCGTGCAAGTCTTACGCGCTTACCCCGACCAATGAG	240
Db	181	CTTCGTGACCAACCGCTGACCTTACCGGCGTGCAAGTCTTACGCGCTTACCCCGACCAATGAG	240
QY	241	CAGCAGACATTTCTTCAATGCGGCCATGCGCCGAAGCTTACGTCCAGAGGCGACATCTTC	300
Db	241	CAGCAGACATTTCTTCAATGCGGCCATGCGCCGAAGCTTACGTCCAGAGGCGACATCTTC	300
QY	301	TTTCAAGAGACGACGGCAACTACAAGACCGCGCCGAGGTAAGTTTCAGAGGGCGACACCTTG	360
Db	301	TTTCAAGAGACGACGGCAACTACAAGACCGCGCCGAGGTAAGTTTCAGAGGGCGACACCTTG	360
QY	361	GTGAACCGCATCGAGCTGAAAGGGCATGCACTTCAAGAGGAGCGGCAACATCTGGGGCAC	420
Db	361	GTGAACCGCATCGAGCTGAAAGGGCATGCACTTCAAGAGGAGCGGCAACATCTGGGGCAC	420
QY	421	AAGCTGAGGTAACAATTACAACGCCACAACGTCTTATCATATGCGCCACAAAGCAGAAAGAC	480
Db	421	AAGCTGAGGTAACAATTACAACGCCACAACGTCTTATCATATGCGCCACAAAGCAGAAAGAC	480
QY	481	GGCATCAAGTGMACTTCAAGATCGGCACACAATCGAGGACGGCACGCTGACGCTGGCC	540
Db	481	GGCATCAAGTGMACTTCAAGATCGGCACACAATCGAGGACGGCACGCTGACGCTGGCC	540
QY	541	GACCACTACACAGAGAAACACCCCATCGGCGACGGCCCGTGTCTGCTGCCGACACAC	600
Db	541	GACCACTACACAGAGAAACACCCCATCGGCGACGGCCCGTGTCTGCTGCCGACACAC	600
QY	601	TACCTGAGACACCACTCGCCCTGAGCAAGAACCACAACGAAAGGCGCATTCACATGTC	660
Db	601	TACCTGAGACACCACTCGCCCTGAGCAAGAACCACAACGAAAGGCGCATTCACATGTC	660
QY	661	CTGCTGAGAGTTCTGTACCGCGCGCGGAGTCACTCTTGCGCATGAGCGAGCTGTACAAGAG	720
Db	661	CTGCTGAGAGTTCTGTACCGCGCGCGGAGTCACTCTTGCGCATGAGCGAGCTGTACAAGAG	720
QY	721	CTTACGCAATGCTTCCCGCGGAGGTGAGAGAGAGATATATGAGCAGCTGCCCATGTCT	780
Db	721	CTTACGCAATGCTTCCCGCGGAGGTGAGAGAGAGATATATGAGCAGCTGCCCATGTCT	780
QY	781	TGTGCGCAGAGAGCGAGTGAACCTGTACCTTCAACCTGTGCTTCTGTAGATCAAT	840
Db	781	TGTGCGCAGAGAGCGGAGTGAACCTGTACCTTCAACCTGTGCTTCTGTAGATCAAT	840
QY	841	GTTGTAGATGC	850
Db	841	GTTGTAGATGC	850

RESULT 2
US-09-191-233-2
; Sequence 2, Application US/09191233A
; GENERAL INFORMATION:

QY 541 GACCACTACAGAGAGAACCCCATTCGGGACGGCCCGTGTGTCGCCCAACAC 600
Db 541 GACCACTACAGAGAGAACCCCATTCGGGACGGCCCGTGTGTCGCCCAACAC 600
QY 601 TACCTGAGACACCGAGCCCGCTGAGCAAAAGCCCAAGAGAGGGGATCATGTGTC 660
Db 601 TACCTGAGACACCGAGCCCGCTGAGCAAAAGCCCAAGAGAGGGGATCATGTGTC 660
QY 661 CTGCTGAGGTTGTCGACCGCCCGGAGTCACTTCGGCATGAGACAGTGTCAAGAAG 720
Db 661 CTGCTGAGGTTGTCGACCGCCCGGAGTCACTTCGGCATGAGACAGTGTCAAGAAG 720
QY 721 CTTAAGCATGCTTCCCGCCGAGGTGTGAGAGACAGAGATGATGACGCTGCCATGTCT 780
Db 721 CTTAAGCATGCTTCCCGCCGAGGTGTGAGAGACAGAGATGATGACGCTGCCATGTCT 780
QY 781 TGTGCCAGAGAGAGCGGATGAGACCGTCACTTCGACGCTGTGCTGTAGATCAAT 840
Db 781 TGTGCCAGAGAGAGCGGATGAGACCGTCACTTCGACGCTGTGCTGTAGATCAAT 840
QY 841 GTGTAGATGC 850
Db 841 GTGTAGATGC 850

RESULT 4
US-09-365-089-2

; Sequence 2, Application US/09365089
; GENERAL INFORMATION:
; APPLICANT: Kain, Steve
; APPLICANT: Li, Xiangliang
; TITLE OF INVENTION: Rapidly Degrading GFP-Fusion Proteins and Methods
; FILE REFERENCE: D6100CIP/D1
; CURRENT APPLICATION NUMBER: US/09/365,089
; CURRENT FILING DATE: 1999-07-30
; EARLIER APPLICATION NUMBER: US 09/191,233
; EARLIER FILING DATE: 1998-11-13
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 2
; LENGTH: 850
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: DNA sequence which encodes the EGFP-MODC422-461
; US-09-365-089-2

Query Match 100.0%; Score 850; DB 19; Length 850;
Best Local Similarity 100.0%; Pred. No. 2.5e-134;
Matches 850; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTGAGCAAGGCGAGAGCTGTTCAACCGGGGTGTGCCATCTGTGTCAGCTGAGC 60
Db 1 ATGTGAGCAAGGCGAGAGCTGTTCAACCGGGGTGTGCCATCTGTGTCAGCTGAGC 60
QY 61 GGCACAGTAAACGGCCACAAGTTCAAGCTGTCCGGCGAGGGCGAGGGGATGCCACTAC 120
Db 61 GGCACAGTAAACGGCCACAAGTTCAAGCTGTCCGGCGAGGGCGAGGGGATGCCACTAC 120
QY 121 GGCAGCTGACCCCTGAAGTTGATCTGACCAACCGGCAAGCTGCCCTGTGCCCCAAC 180
Db 121 GGCAGCTGACCCCTGAAGTTGATCTGACCAACCGGCAAGCTGCCCTGTGCCCCAAC 180
QY 181 CTGTGTGACCAACCTGACCTACGAGCGTGTGTCAGCCGCTACCCCGACCATGAAG 240
Db 181 CTGTGTGACCAACCTGACCTACGAGCGTGTGTCAGCCGCTACCCCGACCATGAAG 240
QY 241 CAGCAGCATCTTCTCAAGTCCGCGATGCCGAGAGGCTACGTCAGAGAGCGACCATCTTC 300
Db 241 CAGCAGCATCTTCTCAAGTCCGCGATGCCGAGAGGCTACGTCAGAGAGCGACCATCTTC 300

QY 301 TTCAAGAGAGAGCACTACAGAACCCCGCCGAGGTGAAGTTGAGGGGACACCTTG 360
Db 301 TTCAAGAGAGAGCACTACAGAACCCCGCCGAGGTGAAGTTGAGGGGACACCTTG 360
QY 361 GTGAACCCGATTCGAGCTAAGGGCATGCACTTCAAGAGAGAGCGACATCTGTGGGAC 420
Db 361 GTGAACCCGATTCGAGCTAAGGGCATGCACTTCAAGAGAGAGCGACATCTGTGGGAC 420
QY 421 AAGCTGAGGTCAACCTACAGACCGCAACGCTATATCATGGCCGCAAGCAAGAAC 480
Db 421 AAGCTGAGGTCAACCTACAGACCGCAACGCTATATCATGGCCGCAAGCAAGAAC 480
QY 481 GGCATCAAGGTGAACCTTCAAGATCCGACCAACATCGAGAGCGGAGGTCAGCTGCC 540
Db 481 GGCATCAAGGTGAACCTTCAAGATCCGACCAACATCGAGAGCGGAGGTCAGCTGCC 540
QY 541 GACCACTACAGAGAGACCCCGATGAGGAGCGGCCCGCTGTGCTGCCCAACAC 600
Db 541 GACCACTACAGAGAGACCCCGATGAGGAGCGGCCCGCTGTGCTGCCCAACAC 600
QY 601 TACCTGAGACACCGAGTCCGCTGAGCAAGACCCCAAGAGAGCGGATCAATGCTC 660
Db 601 TACCTGAGACACCGAGTCCGCTGAGCAAGACCCCAAGAGAGCGGATCAATGCTC 660
QY 661 CTGCTGAGGTTGTCGACCGCCCGGATCACTCTCGGCATGAGCAGCTGTACAGAGAG 720
Db 661 CTGCTGAGGTTGTCGACCGCCCGGATCACTCTCGGCATGAGCAGCTGTACAGAGAG 720
QY 721 CTTAAGCATGCTTCCCGCCGAGGTGTGAGAGAGAGATGATGACGAGCTGCCATGTCT 780
Db 721 CTTAAGCATGCTTCCCGCCGAGGTGTGAGAGAGAGATGATGACGAGCTGCCATGTCT 780
QY 781 TGTGCCAGAGAGAGCGGATGAGACCGTCACTTCGACGCTGTGCTGTAGATCAAT 840
Db 781 TGTGCCAGAGAGAGCGGATGAGACCGTCACTTCGACGCTGTGCTGTAGATCAAT 840
QY 841 GTGTAGATGC 850
Db 841 GTGTAGATGC 850

RESULT 5
US-09-931-232-2

; Sequence 2, Application US/09931232
; GENERAL INFORMATION:
; APPLICANT: Kain, Steve
; APPLICANT: Li, Xiangliang
; TITLE OF INVENTION: Rapidly Degrading GFP-Fusion Proteins
; FILE REFERENCE: CLON075CON
; CURRENT APPLICATION NUMBER: US/09/931,232
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/364,946
; PRIOR FILING DATE: 1999-07-30
; PRIOR APPLICATION NUMBER: 09/191,233
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: 09/062,102
; PRIOR FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: 60/060,855
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 850
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Sequence of the EGFP-MODC422_461 fusion
; OTHER INFORMATION: protein.
; US-09-931-232-2

Query Match 100.0%; Score 850; DB 38; Length 850;
Best Local Similarity 100.0%; Pred. No. 2.5e-134;

Matches 850; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTGAGCAAGGGGAGAGAGCTGTTCACCGGGGTGTGCCATCTGTGAGCTGAGC 60
Db 1 ATGTGAGCAAGGGGAGAGAGCTGTTCACCGGGGTGTGCCATCTGTGAGCTGAGC 60
QY 61 GGGGACGTAAAGGCGCAAGATTCAAGCGGTGTCCGGGAGAGGCGAGCGGCGATCCACTAC 120
Db 61 GGGGACGTAAAGGCGCAAGATTCAAGCGGTGTCCGGGAGAGGCGAGCGGCGATCCACTAC 120
QY 121 GGGGACGTAAAGGCGCAAGATTCAAGCGGTGTCCGGGAGAGGCGAGCGGCGATCCACTAC 180
Db 121 GGGGACGTAAAGGCGCAAGATTCAAGCGGTGTCCGGGAGAGGCGAGCGGCGATCCACTAC 180
QY 181 CTGTGACCAACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTAC 240
Db 181 CTGTGACCAACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTAC 240
QY 241 CAGGACGATCTTCTTCAAGTGTCCGCGCATGCTTCAAGTGTCCGCGCATGCTTCAAGTGTCC 300
Db 241 CAGGACGATCTTCTTCAAGTGTCCGCGCATGCTTCAAGTGTCCGCGCATGCTTCAAGTGTCC 300
QY 301 TTCAAGGACGAGCGGCACTCAAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
Db 301 TTCAAGGACGAGCGGCACTCAAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
QY 361 GTGAACCGCATGAGCTGAGAGGCGCATGCTTCAAGGAGGAGCGGCACTCTGTGGGCGAC 420
Db 361 GTGAACCGCATGAGCTGAGAGGCGCATGCTTCAAGGAGGAGCGGCACTCTGTGGGCGAC 420
QY 421 AAGGTGAGTCACTCAAGACCGCGCGCATGCTTCAAGTGTCCGCGCATGCTTCAAGTGTCC 480
Db 421 AAGGTGAGTCACTCAAGACCGCGCGCATGCTTCAAGTGTCCGCGCATGCTTCAAGTGTCC 480
QY 481 GGCATCAAGTGAATCTCAAGTCTCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540
Db 481 GGCATCAAGTGAATCTCAAGTCTCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540
QY 541 GACGACCTACGAGCAAGACCG 600
Db 541 GACGACCTACGAGCAAGACCG 600
QY 601 TACCTGAGCAAGCGGCACTCAAGTCTCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 660
Db 601 TACCTGAGCAAGCGGCACTCAAGTCTCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 660
QY 661 CTGTGAGTCTGAGCG 720
Db 661 CTGTGAGTCTGAGCG 720
QY 721 CTTAGCCATGAGCTTCCG 780
Db 721 CTTAGCCATGAGCTTCCG 780
QY 781 TGTGACCGAGAGAGCGGAGTGAACCGTCAACCGTCAACCGTCAACCGTCAACCGTCAACCG 840
Db 781 TGTGACCGAGAGAGCGGAGTGAACCGTCAACCGTCAACCGTCAACCGTCAACCGTCAACCG 840
QY 841 GTGTAGATGC 850
Db 841 GTGTAGATGC 850

RESULT 6
PCT-US02-17452-87
; Sequence 87, Application PC/TUS0217452
; GENERAL INFORMATION:
; APPLICANT: CHROMOS MOLECULAR SYSTEMS, INC.
; APPLICANT: Perkins, Edward
; APPLICANT: Perez, Carl
; APPLICANT: Lindenbaum, Michael
; APPLICANT: Greene, Amy
; APPLICANT: Leung, Josephine

APPLICANT: Fleming, Elena
APPLICANT: Stewart, Sandra
APPLICANT: Shellard, Joan
TITLE OR INVENTION: CHROMOSOME-BASED PLATFORMS
FILE REFERENCE: 24601-420PC
CURRENT APPLICATION NUMBER: PCT/US02/17452
PRIOR FILING DATE: 2002-05-30
PRIOR APPLICATION NUMBER: 60/294,758
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: 60/366,891
PRIOR FILING DATE: 2002-03-21
NUMBER OF SEQ ID NOS: 129
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 87
LENGTH: 4862
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: pD2egFP-1N plasmid from Clontech
PCT-US02-17452-87

Query Match 100.0%; Score 850; DB 1; Length 4862;
Best Local Similarity 100.0%; Pred. No. 2.2e-134;
Matches 850; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTGAGCAAGGGGAGAGAGCTGTTCACCGGGGTGTGCCATCTGTGAGCTGAGC 60
Db 679 ATGTGAGCAAGGGGAGAGAGCTGTTCACCGGGGTGTGCCATCTGTGAGCTGAGC 738
QY 61 GGGGACGTAAAGGCGCAAGATTCAAGCGGTGTCCGGGAGAGGCGAGCGGCGATCCACTAC 120
Db 739 GGGGACGTAAAGGCGCAAGATTCAAGCGGTGTCCGGGAGAGGCGAGCGGCGATCCACTAC 798
QY 121 GGGGACGTAAAGGCGCAAGATTCAAGCGGTGTCCGGGAGAGGCGAGCGGCGATCCACTAC 180
Db 799 GGGGACGTAAAGGCGCAAGATTCAAGCGGTGTCCGGGAGAGGCGAGCGGCGATCCACTAC 858
QY 181 CTGTGACCAACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTAC 240
Db 859 CTGTGACCAACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTAC 918
QY 241 CAGGACGATCTTCTTCAAGTGTCCGCGCATGCTTCAAGTGTCCGCGCATGCTTCAAGTGTCC 300
Db 919 CAGGACGATCTTCTTCAAGTGTCCGCGCATGCTTCAAGTGTCCGCGCATGCTTCAAGTGTCC 978
QY 301 TTCAAGGACGAGCGGCACTCAAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
Db 979 TTCAAGGACGAGCGGCACTCAAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1038
QY 361 GTGAACCGCATGAGCTGAGAGGCGCATGCTTCAAGGAGGAGCGGCACTCTGTGGGCGAC 420
Db 1039 GTGAACCGCATGAGCTGAGAGGCGCATGCTTCAAGGAGGAGCGGCACTCTGTGGGCGAC 1098
QY 421 AAGGTGAGTCACTCAAGACCGCGCGCATGCTTCAAGTGTCCGCGCATGCTTCAAGTGTCC 480
Db 1099 AAGGTGAGTCACTCAAGACCGCGCGCATGCTTCAAGTGTCCGCGCATGCTTCAAGTGTCC 1158
QY 481 GGCATCAAGTGAATCTCAAGTCTCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540
Db 1159 GGCATCAAGTGAATCTCAAGTCTCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1218
QY 541 GACCACTACGAGCAAGACCG 600
Db 1219 GACCACTACGAGCAAGACCG 1278
QY 601 TACCTGAGCAAGCGGCACTCAAGTCTCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 660
Db 1279 TACCTGAGCAAGCGGCACTCAAGTCTCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1338
QY 661 CTGTGAGTCTGAGCG 720
Db 1339 CTGTGAGTCTGAGCG 1398

QY 721 CTTAGCATGCTTCCCGCGAGAGTGAGAGCAGATGATGCAAGCTGCCATGCT 780
DB 1399 CTTAGCATGCTTCCCGCGAGAGTGAGAGCAGATGATGCAAGCTGCCATGCT 1458
QY 781 TGTGCCAGAGAGGCGGATGAGACCTGCAAGCTGCTGCTTCTGTGATCAAT 840
DB 1459 TGTGCCAGAGAGGCGGATGAGACCTGCAAGCTGCTGCTTCTGTGATCAAT 1518
QY 841 GTGTAGATGC 850
DB 1519 GTGTAGATGC 1528

RESULT 7
US-10-161-403-87
; Sequence 87, Application US/10161403
; GENERAL INFORMATION:
; APPLICANT: Perkins, Edward
; APPLICANT: Perez, Carl
; APPLICANT: Lindenbaum, Michael
; APPLICANT: Greene, Amy
; APPLICANT: Leung, Josephine
; APPLICANT: Fleming, Elena
; APPLICANT: Stewart, Sandra
; APPLICANT: Shellard, Joan
; TITLE OF INVENTION: CHROMOSOME-BASED PLATFORMS
; FILE REFERENCE: 24601-420
; CURRENT APPLICATION NUMBER: US/10/161,403
; CURRENT FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: 60/294,758
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: 60/366,891
; PRIOR FILING DATE: 2002-03-21
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 87
; LENGTH: 4862
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: pd2egfp-IN plasmid from Clontech
US-10-161-403-87

Query Match 100.0%; Score 850; DB 46; Length 4862;
Best Local Similarity 100.0%; Pred. No. 2,2e-134;
Matches 850; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTGAGCAAGGCGAGAGCTGTTCAAGCGGAGTGGCCCATCTGTGAGCTGAC 60
DB 679 ATGTGAGCAAGGCGAGAGCTGTTCAAGCGGAGTGGCCCATCTGTGAGCTGAC 738
QY 61 GGCACGTAACCGCCCAAGTTAGCGTGTCCGCGAGAGGCGAGGCGATGACACTAC 120
DB 739 GGCACGTAACCGCCCAAGTTAGCGTGTCCGCGAGAGGCGAGGCGATGACACTAC 798
QY 121 GGCAGCTGACCCCTGAAGTTATCTGCAACCCGCAAGCTGCCCTGAGCCCAACC 180
DB 799 GGCAGCTGACCCCTGAAGTTATCTGCAACCCGCAAGCTGCCCTGAGCCCAACC 858
QY 181 CTGCTGACCAAGCTGACCTGACGAGGTGAGTTCAGCGGCTGACCCCGACCAATAG 240
DB 859 CTGCTGACCAAGCTGACCTGACGAGGTGAGTTCAGCGGCTGACCCCGACCAATAG 918
QY 241 CAGACGACCTTCTTCAAGTCCGATGCGCCGAAAGCTACGTCAGAGAGCAGCATCTTC 300
DB 919 CAGACGACCTTCTTCAAGTCCGATGCGCCGAAAGCTACGTCAGAGAGCAGCATCTTC 978
QY 301 TTCAAGACGACGAGCACTCAAGACCCGCGGCGAGAGTTCAGAGGCGAGCACTTCG 360
DB 979 TTCAAGACGACGAGCACTCAAGACCCGCGGCGAGAGTTCAGAGGCGAGCACTTCG 1038
QY 361 GTGAACCGCATGAGCTGGAAGGCGATGACTTCAAGAGAGAGCGGCAATCTGTGGGAC 420

DB 1039 GTGAACCGCATGAGCTGGAAGGCGATGAGCTTCAAGAGAGAGCGGCAATCTGTGGGAC 1098
QY 421 AAGCTGAGGTCACTATCAAGCCCAAGCTCTATATCATGAGCCGACAGAGAGAG 480
DB 1099 AAGCTGAGGTCACTATCAAGCCCAAGCTCTATATCATGAGCCGACAGAGAGAG 1158
QY 481 GGCATCAAGGTGAAGCTTCAAGATCCGACCAAGCTGAGAGAGGCGAGCGTCACTG 540
DB 1159 GGCATCAAGGTGAAGCTTCAAGATCCGACCAAGCTGAGAGAGGCGAGCGTCACTG 1218
QY 541 GACCATCAAGCAAGCAAGCAAGCCGATGAGAGAGGCGGCTGCTGAGCCGACCAAG 600
DB 1219 GACCATCAAGCAAGCAAGCAAGCCGATGAGAGAGGCGGCTGCTGAGCCGACCAAG 1278
QY 601 TACCTGAGCAAGCAAGCTGAGCCGATGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
DB 1279 TACCTGAGCAAGCAAGCTGAGCCGATGAGAGAGAGAGAGAGAGAGAGAGAGAG 1338
QY 661 CTGCTGAGGTGAGACCGCGCGGAGTCACTCTCGGATAGAGAGAGAGAGAGAG 720
DB 1339 CTGCTGAGGTGAGACCGCGCGGAGTCACTCTCGGATAGAGAGAGAGAGAGAG 1398
QY 721 CTTAGCATGCTTCCCGCGAGAGTGAGAGAGAGATGAGAGAGAGAGAGAGAGAG 780
DB 1399 CTTAGCATGCTTCCCGCGAGAGTGAGAGAGAGAGATGAGAGAGAGAGAGAGAG 1458
QY 781 TGTGCCAGAGAGGCGGATGAGACCTGCAAGCTGCTGCTGTGATCAAT 840
DB 1459 TGTGCCAGAGAGGCGGATGAGACCTGCAAGCTGCTGCTGTGATCAAT 1518
QY 841 GTGTAGATGC 850
DB 1519 GTGTAGATGC 1528

RESULT 8
PCT-US99-28451-12
; Sequence 12, Application PC/TUS9928451
; GENERAL INFORMATION:
; APPLICANT: Li, Xiangqiang
; APPLICANT: Zhao, Xiaoning
; APPLICANT: Fang, Yu
; APPLICANT: Jiang, Xin
; APPLICANT: Duong, Tommy
; APPLICANT: Kain, Steve R.
; TITLE OF INVENTION: Cis-Element Reporter Constructs and Uses Thereof
; FILE REFERENCE: D6170PCT
; CURRENT APPLICATION NUMBER: PCT/US99/28451
; CURRENT FILING DATE: 1999-12-01
; EARLIER APPLICATION NUMBER: US 09/206,887
; EARLIER FILING DATE: 1998-12-08
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 12
; LENGTH: 4163
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: synthesized pNF-(B-d2EGFP construct
PCT-US99-28451-12

Query Match 99.5%; Score 846; DB 1; Length 4163;
Best Local Similarity 100.0%; Pred. No. 1.1e-133;
Matches 846; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTGAGCAAGGCGAGAGCTGTTCAAGCGGAGTGGCCCATCTGTGAGCTGAC 60
DB 241 ATGTGAGCAAGGCGAGAGCTGTTCAAGCGGAGTGGCCCATCTGTGAGCTGAC 300
QY 61 GGCAGTAAACCGCCCAAGTTACGCTGTCCGCGAGAGGCGAGGCGATGACACTAC 120
DB 301 GGCAGTAAACCGCCCAAGTTACGCTGTCCGCGAGAGGCGAGGCGATGACACTAC 360
QY 121 GGCAGTAAACCGCCCAAGTTACGCTGTCCGCGAGAGGCGAGGCGATGACACTAC 180

Db 361 GGCAGAGTGAACCTTAAAGTTTATCTGGAACACCGCAAGCTGCGCCGTGCGCCAC 420
Qy 181 CTGCGACCAACCTGACCTTAAAGCGGTGAGTCTTCAACCGCTTACCCGACCACTGAG 240
Db 421 CTGCGACCAACCTGACCTTAAAGCGGTGAGTCTTCAACCGCTTACCCGACCACTGAG 480
Qy 241 CAGACGACTTCTTCAAGTCCGCAATGCGCCGCAAGGCTACGTCAAGAGCGACATCTTC 300
Db 481 CAGACGACTTCTTCAAGTCCGCAATGCGCCGCAAGGCTACGTCAAGAGCGACATCTTC 540
Qy 301 TTCAAGAGCAAGCGCACTCAAGACCCGCGCGCAAGGTGAAGTTGAGGCGCAACCTTG 360
Db 541 TTCAAGAGCAAGCGCACTCAAGACCCGCGCGCAAGGTGAAGTTGAGGCGCAACCTTG 600
Qy 361 GTGAACCGCATGAGCTGAAGGCGATCGACTTCAAGAGAGAGCGGCAATCTCTGGGGAC 420
Db 601 GTGAACCGCATGAGCTGAAGGCGATCGACTTCAAGAGAGAGCGGCAATCTCTGGGGAC 660
Qy 421 AAGCTGAGTCAACTCAAGACCAAGCAACGCTATATCATGCGCGCAAGCAAGAGAAC 480
Db 661 AAGCTGAGTCAACTCAAGACCAAGCAACGCTATATCATGCGCGCAAGCAAGAGAAC 720
Qy 481 GGCATCAAGGTGAACCTTCAAGATCCGCAACAATCGAGAGCGAGCGTCAAGTGGCC 540
Db 721 GGCATCAAGGTGAACCTTCAAGATCCGCAACAATCGAGAGCGAGCGTCAAGTGGCC 780
Qy 541 GACCACTACAGAGCAACACCCCATGCGCGAGCGCGCGCGCGCGCGCGCGCGCAACAC 600
Db 781 GACCACTACAGAGCAACACCCCATGCGCGAGCGCGCGCGCGCGCGCGCGCGCAACAC 840
Qy 601 TACCTGAGCAACCGATCCGCGCTGAGCAAAACCCCAAGAGCGGATCATGCTGC 660
Db 841 TACCTGAGCAACCGATCCGCGCTGAGCAAAACCCCAAGAGCGGATCATGCTGC 900
Qy 661 CTGCTGAGTTCGTAACCG 720
Db 901 CTGCTGAGTTCGTAACCG 960
Qy 721 CTGCTGAGTTCGTAACCG 780
Db 961 CTGCTGAGTTCGTAACCG 1020
Qy 781 TGTGCCAGAGAGCGGATGACCTGACCTGAGCGCTGCTTCTGCTAGATCAAT 840
Db 1021 TGTGCCAGAGAGCGGATGACCTGACCTGAGCGCTGCTTCTGCTAGATCAAT 1080
Qy 841 GTGTAG 846
Db 1081 GTGTAG 1086

RESULT 9
US-09-206-887-12

; Sequence 12, Application US/09206887A
; GENERAL INFORMATION:
; APPLICANT: Li, Xiangqiang
; APPLICANT: Zhao, Xiaoning
; APPLICANT: Pang, Yu
; APPLICANT: Jiang, Xin
; APPLICANT: Duong, Tommy
; APPLICANT: Kahn, Steve R.
; TITLE OF INVENTION: cis-Element Reporter Constructs and Uses Thereof
; FILE REFERENCE: D6170
; CURRENT APPLICATION NUMBER: US/09/206,887A
; CURRENT FILING DATE: 1998-12-08
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 12
; LENGTH: 4163
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: synthesized pNF-(B-d2EGFP construct

US-09-206-887-12

Query Match 99.5%; Score 846; DB 16; Length 4163;
Best Local Similarity 100.0%; Pred. No. 1,1e-133;
Matches 846; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGCTGAGCAAGGCGAGAGCTGTTACCGGGGTGGTGGCCATCTGTGTGACCTGAC 60
Db 241 ATGCTGAGCAAGGCGAGAGCTGTTACCGGGGTGGTGGCCATCTGTGTGACCTGAC 300
Qy 61 GGCAGCGTAAACGGGCAACAGTTCAAGCTGTCCGCGCAAGGCGAGGCGATGCCACTAC 120
Db 301 GGCAGCGTAAACGGGCAACAGTTCAAGCTGTCCGCGCAAGGCGAGGCGATGCCACTAC 360
Qy 121 GGCAGGCTGACCTGAAAGTTTATCTGACCAACCGGCAAGCTGCGCGCTGCTGCGCCAC 180
Db 361 GGCAGGCTGACCTGAAAGTTTATCTGACCAACCGGCAAGCTGCGCGCTGCTGCGCCAC 420
Qy 181 CTGCTGACCAACCTGACCTGACGCGGTGCAAGTCTTCAAGCGCTACCCCGCAACATGAG 240
Db 421 CTGCTGACCAACCTGACCTGACGCGGTGCAAGTCTTCAAGCGCTACCCCGCAACATGAG 480
Qy 241 CAGACGACTTCTTCAAGTCCGCGCATGCGCGCAAGGCTACGTCAAGAGCGGACCATCTTC 300
Db 481 CAGACGACTTCTTCAAGTCCGCGCATGCGCGCAAGGCTACGTCAAGAGCGGACCATCTTC 540
Qy 301 TTCAAGAGCAAGCGGCACTCAAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
Db 541 TTCAAGAGCAAGCGGCACTCAAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
Qy 361 GTGAACCGCATGAGCTGAAGGCGATGACTTCAAGAGAGAGCGGCAATCTCTGGGGAC 420
Db 601 GTGAACCGCATGAGCTGAAGGCGATGACTTCAAGAGAGAGCGGCAATCTCTGGGGAC 660
Qy 421 AAGCTGAGTCAACTCAAGACCAAGCAACGCTATATCATGCGCGCAAGCAAGAGAAC 480
Db 661 AAGCTGAGTCAACTCAAGACCAAGCAACGCTATATCATGCGCGCAAGCAAGAGAAC 720
Qy 481 GGCATCAAGGTGAACCTTCAAGATCCGCGCAACAATCGAGAGCGGACCGTGGCGCC 540
Db 721 GGCATCAAGGTGAACCTTCAAGATCCGCGCAACAATCGAGAGCGGACCGTGGCGCC 780
Qy 541 GACCACTACAGAGCAACACCCCATGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
Db 781 GACCACTACAGAGCAACACCCCATGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 840
Qy 601 TACCTGAGCAACCGATCCGCGCTGAGCAAAACCCCAAGAGCGGATCATGCTGC 660
Db 841 TACCTGAGCAACCGATCCGCGCTGAGCAAAACCCCAAGAGCGGATCATGCTGC 900
Qy 661 CTGCTGAGTTCGTAACCG 720
Db 901 CTGCTGAGTTCGTAACCG 960
Qy 721 CTGCTGAGTTCGTAACCG 780
Db 961 CTGCTGAGTTCGTAACCG 1020
Qy 781 TGTGCCAGAGAGCGGATGACCTGACCTGAGCGCTGCTTCTGCTAGATCAAT 840
Db 1021 TGTGCCAGAGAGCGGATGACCTGACCTGAGCGCTGCTTCTGCTAGATCAAT 1080
Qy 841 GTGTAG 846
Db 1081 GTGTAG 1086

RESULT 10

PCT-US99-28451-13
; Sequence 13, Application PC/TUS9928451
; GENERAL INFORMATION:
; APPLICANT: Li, Xiangqiang
; APPLICANT: Zhao, Xiaoning


```

1  APPLICANT: Fang, Yu
2  APPLICANT: Jiang, Xin
3  APPLICANT: Duong, Tommy
4  APPLICANT: Kain, Steve R.
5  TITLE OF INVENTION: Cis-Element Reporter Constructs and Uses Thereof
6  FILE REFERENCE: D6170DPT
7  CURRENT APPLICATION NUMBER: PCT/US99/28451
8  CURRENT FILING DATE: 1999-12-01
9  EARLIER APPLICATION NUMBER: US 09/206,887
10 EARLIER FILING DATE: 1998-12-08
11 NUMBER OF SEQ ID NOS: 14
12 SEQ ID NO 13
13     LENGTH: 4300
14     TYPE: DNA
15     ORGANISM: artificial sequence
16     FEATURE:
17 OTHER INFORMATION: synthesized pCRE5-d2B6FP
18 PCT-US99-28451-13

```

Query Match 99.5%; Score 846; DB 1; Length 4300;
Best Local Similarity 100.0%; Pred. No. 1,1e-13;
Matches 846; Conservative 0; Mismatches 0; Gaps 0;

OY 1 ATGGTGGACAAAGGCGAGAGAGCTCTTACACCGAGGTGGATCCCAATCTGGTGGACCTGGAC 60
Db 378 ATGGTGGACAAAGGCGAGAGAGCTCTTACACCGAGGTGGATCCCAATCTGGTGGACCTGGAC 437

438 GCGCAGCTAAACGGGCCAAGATTCAGCGTCCGGCGAGGCGCGATGCCACTAC 497

D2 498 GGCAGCTGACCCCTGAAGTTCACTGACCAACGCGAAAGCTGCGCCGTGCGCTGGGCCACC 557

558 CTGCTGACCAACCTGACTACCGCGTGCAGTCTTACGCGCTACCCCGACCAATGAAG 617

	Db	618	CAGCACAACCTTCTCAAGTCGCCGCATGCGCCGAAGGCTAAGTCCAGAGAGGCACCACCATCTTC	677
301	TTC	CAAGAGACGACGGACACTTACAAAGACCCCGCGGAGAGGAAAGATTGAGAGGCGGAGAACCCGTA	360	

Db 678 TTCAAGGACGACGGCAACTTACAAGACCACCGCCGACCGAGGTAAGTTGAGGGCGACACCCCTG 737

Db 738 GTGACCCGATCGAGTGAAGGCGATTCGACCTTCAAGAGAGACGGCGACATCTCTGGGCGAC 797

OY 421 AAGCTGGAGTACCACTCAACACGCGACAAAGTCTATATATCATGCGCGACAGCAGAGGAC 480

Db 798 AAGCTGAGTACACTCAACAGGCCACAAGTGTATATCATGTGGCGACACAGCAGAGAAC 857

Dy 481 GGCATCAAGGTAACTTCAAGATCCGCCCAACATCGAGGACGGACGGCTGACGTGCC 540

Db 858 GGCATCAGAAGTAACTTCAAGATCCGCGACAAACATCGAGGACGGCAGCGCTGCAGCTCGCC 917
 QY 541 GACCACTACAGAGAAACACCCCATCGCGGACGGCCCGCTGGCTGTGCCGGAACAC 600

D5 518 TACCCATCAACAGTAAACACCCCATCCGAGACGGCCCCCGTGCCTGCCCCGACACCC 977

QY 601 TACCTGACACCCAGTCCGCCCTGAGCAAAAGCCCCAAAGAAAGCGCATCATGTGTC 660

661 CTGCTGGAATTGTGACCGCGCGCGGAGTCACTCTGGGATGAGCGAGCTGTACAGAAG 720

721 CTTAGGCATGGCTTCCGCCCGAGCTGGAGGAGAGATGATGGCACGCTGCCATGTCT 780

Db	1098	CTTAGCCATGCGTTC	CGCGCGAGGTGGAAGACGAGATGATGGCAGCGTCCCATGTC	1157
QY	781	TTGAGCCCAAGAGACGGGAGATGACACCGTCAACCGTGGTTCGTGCTAAGGATCAAT		840
Db	1158	TGTGCCCAAGAGACGGGATGACACCGTCAACCGTGGTTCGTGCTAAGATCAAT		1217
QY	841	GTTGTAG	846	
Db	1218	GTTGTAG	1223	

RESULT 11
US-09-206-887-13
Sentence 13
url:https://www.fda.gov/oc/foia/US-09-206-887-13

```

: GENERAL INFORMATION:
: APPLICANT: Li, Xianqiang
: APPLICANT: Zhao, Xiaoning
: APPLICANT: Pang, Yu

```

AFFILIANT: Daug, Ah
 APPLICANT: Duong, Tommy
 APPLICANT: Kain, Steve R.
 TITLE OF INVENTION: Cis-Element Reporter Constructs and Uses Thereof

```

1 CURRENT APPLICATION NUMBER: US/09/206,887A
2
3 CURRENT FILING DATE: 1998-12-08
4
5 NUMBER OF SEQ ID NOS: 14
6

```

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Query Match	99.5%	Score 846:	DB 16:	length 4300:
OTHER INFORMATION: synthesized pCRES-d2EGFP				
US-09-206-887-13				

Query Match	99.5%	Score 846	DB 16	Length 4300
Best Local Similarity	100.0%	Pred. No. 1.1e-133		
Matches 846	Conservative 0	Mismatches 0	Indels 0	Gaps 0
1 ATGGTGTGACCAAGGAGGAGACTGTTTACCCAGGGGTGTGTCGCATCTGTCGTGAGACTGGAC 60				

Db 378 ATGCTGACCAAGGCGCCGAGAGCTGTTACCCGGGGTGGTCCCATCTCTGAGC 437

Dy 61 GCGACGTAACCGGCCACAAGTTTCAGCGTGTCCGCGAGGGCGAGGGCGATGCCACTTAC 120

Db 438 GCGACGCTAAACGGCCCAAGTTCAAGCTGTCCGGCGACGAGGGCGATGCCACCTAC 497

D6 498 GGCACGCTGACCCCTGAAATTCACTGCAACCAACCGGCAAGCTGCGCGTGGCCCTGGCCCCAC 557

DY 181 CTCGTGACCAACCTGACCTAAGCGCGTGAGTGGTTCAAGCGCGGTAAACCCGACCAATATGAG 240

DB 558 CTGCGTGCACCCCTGACCTTACGGGCGTGAGTCTTCAGCCCGCTTACCCCGACCCCATATGAG 61

QY 241 CAGCAGCACTTCTTCAAGTCGCCCATGCCCGGAGAGCTACGTCCAGAGACGCACCATCTTC 300

301 TTCAAGGACGACGGCAA CTACAGAGCCCCGCCGAGGTGAAGTTCGAGGGCGACCCCTG 360

QY 361 GTGAACCGCATGAGCTGGAAGGCATCGACTTGAAGGAGGACGGCAACATCTCGGGGAC 420

421 AAGCTGGAGTAACTACAACAGCCACAGCTGTATATCATGGCCGACAAAGCAGAAGAAC 480
OY
798 AAGCTGGAGTAACTACAACAGCCACAGCTGTATATCATGGCCGACAAAGCAGAAGAAC 857
db

QY 481 GGCATCAAGTGAACCTTCAAGATCCGCCACAAATCGAGGACGGACGGTGCAGCTCGCC 540

858 GGCATCAAGTGAACCTTCAAGATCCGCCACAAATCGAGGACGGACGGTGCAGCTCGCC 917

QY 541 GACCACCTACGACGAGAACACCCCATCGGCGACGCGCCCGTGTCTGTCGCCGACCAACAC 600
Db 918 GACCACTACGACGAGAACACCCCATCGGCGACGCGCCCGTGTCTGTCGCCGACCAACAC 977
QY 601 TACTGAGACCCAGTCCGCTGAGCAAGAAGCCCAAGAGAGAGAGAGAGAGAGAGAGAGAG 660
Db 978 TACTGAGACCCAGTCCGCTGAGCAAGAAGCCCAAGAGAGAGAGAGAGAGAGAGAGAGAG 1037
QY 661 CTGCTGAGGTTCTGTAACCGCGCGCGCGGATCACTCTCGGATGAGAGAGAGAGAGAGAG 720
Db 1038 CTGCTGAGGTTCTGTAACCGCGCGCGGATCACTCTCGGATGAGAGAGAGAGAGAGAG 1097
QY 721 CTAGCCATGCTTCCCGCGAG 780
Db 1098 CTAGCCATGCTTCCCGCGAG 1157
QY 781 TGTGCCAG 840
Db 1158 TGTGCCAG 1217
QY 841 GTGTAG 846
Db 1218 GTGTAG 1223

RESULT 12

US-10-332-733-21

Sequence 21, Application US/10332733
GENERAL INFORMATION:
APPLICANT: Margarette Odenhal and Diana Jung
TITLE OF INVENTION: Gene Expression, Genome Alteration And Reporter Expression
TITLE OF INVENTION: In Myofibroblasts And Myofibroblast-like Cells
FILE REFERENCE: 1472/68806
CURRENT APPLICATION NUMBER: US/10/332,733
CURRENT FILING DATE: 2003-06-10
NUMBER OF SEQ ID NOS: 41
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 21
LENGTH: 845
TYPE: DNA
ORGANISM: artificial sequence
FEATURE:
OTHER INFORMATION: Description of artificial sequence: EGFP
US-10-332-733-21

Query Match 99.4%; Score 845; DB 49; Length 845;
Best Local Similarity 100.0%; Pred. No. 1.8e-133;
Matches 845; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGTGAAGCAAGGCGAGAGAGCTGTTCACCGGGGTGGGCCCATCTGCTGAGCTGAGAC 60
Db 1 ATGGTGAAGCAAGGCGAGAGAGCTGTTCACCGGGGTGGGCCCATCTGCTGAGCTGAGAC 60
QY 61 GGCAGCGTAAACGGCCCAAGATTGAGCGTGTCCGCGCAAGGCGAGAGGCGATGCCACTTAC 120
Db 61 GGCAGCGTAAACGGCCCAAGATTGAGCGTGTCCGCGCAAGGCGAGAGGCGATGCCACTTAC 120
QY 121 GGCAGCGTAAACGGCCCAAGATTGAGCGTGTCCGCGCAAGGCGAGAGGCGATGCCACTTAC 180
Db 121 GGCAGCGTAAACGGCCCAAGATTGAGCGTGTCCGCGCAAGGCGAGAGGCGATGCCACTTAC 180
QY 181 CTGCTGAGGTTCTGTAACCGCGCGCGGATCACTCTCGGATGAGAGAGAGAGAGAGAGAG 240
Db 181 CTGCTGAGGTTCTGTAACCGCGCGCGGATCACTCTCGGATGAGAGAGAGAGAGAGAGAG 240
QY 241 CAGGACGATCTTCTCAAGTCCGCGCATGCCGAGGCTTACAGTCCGAGAGAGAGAGAGAGAG 300
Db 241 CAGGACGATCTTCTCAAGTCCGCGCATGCCGAGGCTTACAGTCCGAGAGAGAGAGAGAGAG 300
QY 301 TTCAAG 360
Db 301 TTCAAG 360

QY 361 GTGAACCGCATCGAGCTGAAGAGGAGATGATCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAG 420
Db 361 GTGAACCGCATCGAGCTGAAGAGGAGATGATCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAG 420
QY 421 AAGCTGAGTCACTCACTCAAG 480
Db 421 AAGCTGAGTCACTCACTCAAG 480
QY 481 GGCATCAAGGTGAATCTTCAAGATCCGCGCAACATCGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
Db 481 GGCATCAAGGTGAATCTTCAAGATCCGCGCAACATCGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
QY 541 GACCACTACGACGAGAACACCCCATGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
Db 541 GACCACTACGACGAGAACACCCCATGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
QY 601 TACTGAGACCCAGTCCGCTGAGCAAGAAGCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
Db 601 TACTGAGACCCAGTCCGCTGAGCAAGAAGCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
QY 661 CTGCTGAGGTTCTGTAACCGCGCGCGGATCACTCTCGGATGAGAGAGAGAGAGAGAGAGAG 720
Db 661 CTGCTGAGGTTCTGTAACCGCGCGCGGATCACTCTCGGATGAGAGAGAGAGAGAGAGAGAG 720
QY 721 CTAGCCATGCTTCCCGCGAG 780
Db 721 CTAGCCATGCTTCCCGCGAG 780
QY 781 TGTGCCAG 840
Db 781 TGTGCCAG 840
QY 841 GTGTAG 845
Db 841 GTGTAG 845

RESULT 13

PCT-US03-28111-3

Sequence 3, Application PC/TUS0328111
GENERAL INFORMATION:
APPLICANT: Bear, et al.
TITLE OF INVENTION: Lethal Vectors, Related Reagents, and Methods of
TITLE OF INVENTION: Use thereof
FILE REFERENCE: 0492611-0265
CURRENT APPLICATION NUMBER: PCT/US03/28111
CURRENT FILING DATE: 2003-09-05
NUMBER OF SEQ ID NOS: 39
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 3
LENGTH: 6748
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Plasmid
PCT-US03-28111-3

Query Match 84.9%; Score 721.8; DB 1; Length 6748;
Best Local Similarity 99.0%; Pred. No. 1.1e-112;
Matches 726; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ATGGTGAAGCAAGGCGAGAGAGCTGTTCACCGGGGTGGGCCCATCTGCTGAGCTGAGAC 60
Db 2778 ATGGTGAAGCAAGGCGAGAGAGCTGTTCACCGGGGTGGGCCCATCTGCTGAGCTGAGAC 2837
QY 61 GGCAGCGTAAACGGCCCAAGATTGAGCGTGTCCGCGCAAGGCGAGAGGCGATGCCACTTAC 120
Db 2838 GGCAGCGTAAACGGCCCAAGATTGAGCGTGTCCGCGCAAGGCGAGAGGCGATGCCACTTAC 2897
QY 121 GGCAGCGTAAACGGCCCAAGATTGAGCGTGTCCGCGCAAGGCGAGAGGCGATGCCACTTAC 180
Db 121 GGCAGCGTAAACGGCCCAAGATTGAGCGTGTCCGCGCAAGGCGAGAGGCGATGCCACTTAC 180

US-10-152-040-28

Query Match 84.8%; Score 721.2; DB 46; Length 10417;
Best Local Similarity 97.6%; Pred. No. 1.3e-112;
Matches 732; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

```
QY 1 ATGTGAGCAAGGCGAGGAGCTGTTCAACGGGGGTGTGCCCATCTGTGTGAGCTGAGC 60
DB 743 ATGTGAGCAAGGCGAGGAGCTGTTCAACGGGGGTGTGCCCATCTGTGTGAGCTGAGC 802
QY 61 GGGGAGGTAAAGGCGCAAGTTCAAGCTGTCCGGGAGGGGCGATGCCACTTAC 120
DB 803 GGGGAGGTAAAGGCGCAAGTTCAAGCTGTCCGGGAGGGGCGATGCCACTTAC 862
QY 121 GCGAAGCTGACCTGAAGTTTCATCTSCACACCGGGCAAGCTGCCCTGCGCCACC 180
DB 863 GCGAAGCTGACCTGAAGTTTCATCTSCACACCGGGCAAGCTGCCCTGCGCCACC 922
QY 181 CTGTGACCAACCTGACTGACGGCGGTGACAGTGTTCAGCCGCTACCCCGACACATGAAG 240
DB 923 CTGTGACCAACCTGACTGACGGCGGTGACAGTGTTCAGCCGCTACCCCGACACATGAAG 982
QY 241 CAGCAAGATTCTTCAAGTCCGCGCATGCCCGAAGGCTAGTCCAGGAGCGCACATCTTC 300
DB 983 CAGCAAGATTCTTCAAGTCCGCGCATGCCCGAAGGCTAGTCCAGGAGCGCACATCTTC 1042
QY 301 TTCAAAGACGAGCGCACTACAAAGACCCGCGCGAGGTGAAGTTGAGGGCGACACCTG 360
DB 1043 TTCAAAGACGAGCGCACTACAAAGACCCGCGCGAGGTGAAGTTGAGGGCGACACCTG 1102
QY 361 GTGAACCGCATGAGCTGAAGGGCATTCGACTTCAAGAGAGAGCGCAACATCTTGGGGCAC 420
DB 1103 GTGAACCGCATGAGCTGAAGGGCATTCGACTTCAAGAGAGAGCGCAACATCTTGGGGCAC 1162
QY 421 AAGCTGAGTACACATCAACAAGCCCAAGTCTAATATGAGCCGCAAGAGAAAGAC 480
DB 1163 AAGCTGAGTACACATCAACAAGCCCAAGTCTAATATGAGCCGCAAGAGAAAGAC 1222
QY 481 GGCATCAAGTGAACCTTCAAGATCCGCCCAACATCGAGAGAGCGCAGCTGACCTGCC 540
DB 1223 GGCATCAAGTGAACCTTCAAGATCCGCCCAACATCGAGAGAGCGCAGCTGACCTGCC 1282
QY 541 GACCACTACAGAGAAACACCCCATCGGCGACGGCCCGTGTCTGCTCGACAAACAC 600
DB 1283 GACCACTACAGAGAAACACCCCATCGGCGACGGCCCGTGTCTGCTCGACAAACAC 1342
QY 601 TACCTGAGGACCGAGTCCGCGCTGAGCAAGAAGCCCAAGAGAGCGCATCAATGTTC 660
DB 1343 TACCTGAGGACCGAGTCCGCGCTGAGCAAGAAGCCCAAGAGAGCGCATCAATGTTC 1402
QY 661 CTGCTGAGATTGATGACCGCGCGGGATCACTCTGCGATGAGACGAGCTGTACAAGAG 720
DB 1403 CTGCTGAGATTGATGACCGCGCGGGATCACTCTGCGATGAGACGAGCTGTACAAGAG 1462
QY 721 CTTAGCCATGGCTTCCCGCGGAGGTGAG 750
DB 1463 CTTAGCCATGGCTTCTGTAAAGGAAGTTAAG 1492
```

Search completed: May 15, 2004, 06:21:47
Job time : 3577 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2004, 03:30:36 ; Search time 170 Seconds
(without alignments)
5764.986 Million cell updates/sec

Title: US-09-931-232-2

Perfect score: 850

Sequence: 1 atgtgtgagcagggcgagga.....tagatcaatgtgtatgac 850

Scoring table: IDENTITY_NUC

Searched: 1660298 seqs, 576498597 residues

Total number of hits satisfying chosen parameters: 3320596

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Pending Patents_NA_New.*
1: /cgn2_6/prodata/1/pna/PCT_NEW_COMB.seq.*
2: /cgn2_6/prodata/1/pna/US06_NEW_COMB.seq.*
3: /cgn2_6/prodata/1/pna/US07_NEW_COMB.seq.*
4: /cgn2_6/prodata/1/pna/US08_NEW_COMB.seq.*
5: /cgn2_6/prodata/1/pna/US09_NEW_COMB.seq.*
6: /cgn2_6/prodata/1/pna/US10_NEW_COMB.seq.*
7: /cgn2_6/prodata/1/pna/US60_NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	850	100.0	4336	US-10-679-191A-1	Sequence 1, Appl
2	719.2	84.6	14083	US-10-789-400-3	Sequence 3, Appl
3	718.6	84.5	1190	US-10-811-081-98	Sequence 98, Appl
4	718.6	84.5	4944	US-10-826-523-55	Sequence 55, Appl
5	718.6	84.5	4952	US-10-826-523-52	Sequence 52, Appl
6	718.6	84.5	5451	US-10-485-281-6	Sequence 6, Appl
7	718.6	84.5	9013	US-10-826-523-50	Sequence 50, Appl
8	718.6	84.5	5759	US-09-954-483B-14	Sequence 14, Appl
9	717.4	84.4	1704	US-10-488-758-8	Sequence 8, Appl
10	717.4	84.4	5041	PCT-US04-06378-6	Sequence 6, Appl
11	717.4	84.4	5041	US-10-790-455-6	Sequence 5, Appl
12	717.4	84.4	8633	US-10-475-962-54	Sequence 54, Appl
13	717.4	84.4	717	US-10-785-862-9	Sequence 9, Appl
14	717.4	84.4	8591	US-10-672-764A-42	Sequence 42, Appl
15	717.4	84.4	10089	US-10-672-764A-53	Sequence 53, Appl
16	717.4	84.4	38203	US-10-672-764A-55	Sequence 55, Appl
17	717.4	84.4	38203	US-10-672-764A-56	Sequence 56, Appl
18	714	84.0	765	US-10-129-207B-7	Sequence 7, Appl
19	711	83.6	720	US-10-311-030-10	Sequence 10, Appl
20	707.4	83.2	4941	US-10-826-523-53	Sequence 53, Appl
21	707.4	83.2	4943	US-10-826-523-54	Sequence 54, Appl
22	705.8	83.0	9012	US-10-826-523-49	Sequence 49, Appl
23	704.2	82.8	4944	US-10-826-523-56	Sequence 56, Appl
24	704.2	82.8	4951	US-10-826-523-51	Sequence 51, Appl
25	704.2	82.8	8999	US-10-826-523-48	Sequence 48, Appl
26	703	82.7	1722	US-10-831-622-107	Sequence 107, App

27	589	69.3	600	1	PCT-US04-10982-168	Sequence 168, App
28	584.2	68.7	597	1	PCT-US04-10982-36	Sequence 36, Appl
29	584.2	68.7	597	1	PCT-US04-10982-60	Sequence 60, Appl
30	584.2	68.7	597	1	PCT-US04-10982-164	Sequence 164, Appl
31	584.2	68.7	600	1	PCT-US04-10982-38	Sequence 38, Appl
32	584.2	68.7	600	1	PCT-US04-10982-62	Sequence 62, Appl
33	584.2	68.7	600	1	PCT-US04-10982-166	Sequence 166, Appl
34	582.6	68.5	597	1	PCT-US04-10982-84	Sequence 84, Appl
35	582.6	68.5	597	1	PCT-US04-10982-88	Sequence 88, Appl
36	582.6	68.5	597	1	PCT-US04-10982-152	Sequence 152, Appl
37	582.6	68.5	600	1	PCT-US04-10982-70	Sequence 70, Appl
38	582.6	68.5	600	1	PCT-US04-10982-86	Sequence 86, Appl
39	582.6	68.5	600	1	PCT-US04-10982-154	Sequence 154, Appl
40	581	68.4	597	1	PCT-US04-10982-32	Sequence 32, Appl
41	581	68.4	597	1	PCT-US04-10982-88	Sequence 88, Appl
42	581	68.4	597	1	PCT-US04-10982-96	Sequence 96, Appl
43	581	68.4	597	1	PCT-US04-10982-100	Sequence 100, Appl
44	581	68.4	597	1	PCT-US04-10982-160	Sequence 160, Appl
45	581	68.4	597	1	PCT-US04-10982-174	Sequence 174, App

ALIGNMENTS

RESULT 1

US-10-679-191A-1

Sequence 1, Application US/10679191A

GENERAL INFORMATION:

APPLICANT: Moon, Randall T.

FILE REFERENCE: Dorky, Richard I

CURRENT APPLICATION NUMBER: US/10/679,191A

PRIOR FILING DATE: 2003-10-03

PRIOR FILING DATE: 2002-10-03

NUMBER OF SEQ ID NOS: 4

SOFTWARE: PatentIn version 3.0

SEQ ID NO 1

LENGTH: 4336

TYPE: DNA

ORGANISM: Artificial

FEATURE:

OTHER INFORMATION: plasmid

NAME/KEY: misc feature

LOCATION: (1)..(1)

FEATURE:

NAME/KEY: misc feature

LOCATION: (38)..(243)

OTHER INFORMATION: Topflash enhancer

FEATURE:

NAME/KEY: misc feature

LOCATION: (3485)..(4327)

OTHER INFORMATION: EGFP+pest

US-10-679-191A-1

Query Match 100.0%; Score 850; DB 6; Length 4336;

Best Local Similarity 100.0%; Pred. No. 4.2e-150;

Matches 850; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTGAGCAAGGCGAGGAGCTGTTACCGGGGTGTCCCATCTGTCAGCTGAC 60

DB 3485 ATGTGAGCAAGGCGAGGAGCTGTTACCGGGGTGTCCCATCTGTCAGCTGAC 3544

QY 61 GCGGAGCTTAAAGGCGCAAGTTGACGCTGTCCGGGAGGAGGAGGCGCATGCACTAC 120

DB 3545 GCGGAGCTTAAAGGCGCAAGTTGACGCTGTCCGGGAGGAGGAGGCGCATGCACTAC 3604

QY 121 GGCAAGCTGACCTTGAAGTTCATCTGCACACCGGACACCTCCGTCCTGCGCCACC 180
DB 3605 GGCAAGCTGACCTTGAAGTTCATCTGCACACCGGACACCTCCGTCCTGCGCCACC 3664
QY 181 CTCGTGACCACTTGAAGTTCATCTGCACACCGGACACCTCCGTCCTGCGCCACC 240
DB 3665 CTCGTGACCACTTGAAGTTCATCTGCACACCGGACACCTCCGTCCTGCGCCACC 3724
QY 241 GAGCAGCACTTCTTCAAGTTCGCGCAGGACGCGGACGCGGACGCGGACGCGGAC 300
DB 3725 GAGCAGCACTTCTTCAAGTTCGCGCAGGACGCGGACGCGGACGCGGACGCGGAC 3784
QY 301 TTCAAGAGCAGCAGCACTTCAAGGACGCGGACGCGGACGCGGACGCGGACGCGGAC 360
DB 3785 TTCAAGAGCAGCAGCACTTCAAGGACGCGGACGCGGACGCGGACGCGGACGCGGAC 3844
QY 361 GTGAACCGCATGAGCTGAAGGCGCATCTTCAAGGACGCGGACGCGGACGCGGACGCGGAC 420
DB 3845 GTGAACCGCATGAGCTGAAGGCGCATCTTCAAGGACGCGGACGCGGACGCGGACGCGGAC 3904
QY 421 AAGCTGAGTACCACTTCAAGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGAC 480
DB 3905 AAGCTGAGTACCACTTCAAGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGAC 3964
QY 481 GGCATCAAGTGAAGTCTTCAAGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGAC 540
DB 3965 GGCATCAAGTGAAGTCTTCAAGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGAC 4024
QY 541 GACCACTTCAAGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGAC 600
DB 4025 GACCACTTCAAGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGAC 4084
QY 601 TACCTGAGCAGCAGCTTCAAGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGAC 660
DB 4085 TACCTGAGCAGCAGCTTCAAGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGAC 4144
QY 661 CTGCTGAGTGTGAGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGAC 720
DB 4145 CTGCTGAGTGTGAGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGAC 4204
QY 721 CTTAAGCAGTGTGAGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGAC 780
DB 4205 CTTAAGCAGTGTGAGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGAC 4264
QY 781 TGTGCGCAGAGAGCGGAGTGAACGCTTCAAGGACGCGGACGCGGACGCGGACGCGGACGCGGAC 840
DB 4265 TGTGCGCAGAGAGCGGAGTGAACGCTTCAAGGACGCGGACGCGGACGCGGACGCGGACGCGGAC 4324
QY 841 GTGTAGATGC 850
DB 4325 GTGTAGATGC 4334

RESULT 2
US-10-789-400-3
; Sequence 3, Application US/10789400
; GENERAL INFORMATION:
; APPLICANT: THE GOVERNMENT OF THE UNITED STATES OF AMERICA AS
; APPLICANT: REPRESENTED BY THE SECRETARY OF THE DEPARTMENT OF HEALTH AND
; APPLICANT: HUMAN SERVICES
; APPLICANT: Collins, Peter L.
; APPLICANT: Biacchesi, Stephanie
; APPLICANT: Buchholz, Ursula
; APPLICANT: Skidopoulos, Mario H.
; APPLICANT: Murphy, Brian R.
; TITLE OF INVENTION: RECOMBINANT HUMAN METAPNEUMOVIRUS AND ITS USE
; FILE REFERENCE: 4239-67783
; CURRENT APPLICATION NUMBER: US/10789,400
; CURRENT FILING DATE: 2004-02-27
; PRIOR APPLICATION NUMBER: US 60/451,119
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/478,667

QY 121 GGCAAGCTGACCTTGAAGTTCATCTGCACACCGGACACCTCCGTCCTGCGCCACC 180
DB 3605 GGCAAGCTGACCTTGAAGTTCATCTGCACACCGGACACCTCCGTCCTGCGCCACC 3664
QY 181 CTCGTGACCACTTGAAGTTCATCTGCACACCGGACACCTCCGTCCTGCGCCACC 240
DB 3665 CTCGTGACCACTTGAAGTTCATCTGCACACCGGACACCTCCGTCCTGCGCCACC 3724
QY 241 GAGCAGCACTTCTTCAAGTTCGCGCAGGACGCGGACGCGGACGCGGACGCGGACGCGGAC 300
DB 3725 GAGCAGCACTTCTTCAAGTTCGCGCAGGACGCGGACGCGGACGCGGACGCGGACGCGGAC 3784
QY 301 TTCAAGAGCAGCAGCACTTCAAGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGAC 360
DB 3785 TTCAAGAGCAGCAGCACTTCAAGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGAC 3844
QY 361 GTGAACCGCATGAGCTGAAGGCGCATCTTCAAGGACGCGGACGCGGACGCGGACGCGGACGCGGAC 420
DB 3845 GTGAACCGCATGAGCTGAAGGCGCATCTTCAAGGACGCGGACGCGGACGCGGACGCGGACGCGGAC 3904
QY 421 AAGCTGAGTACCACTTCAAGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGAC 480
DB 3905 AAGCTGAGTACCACTTCAAGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGAC 3964
QY 481 GGCATCAAGTGAAGTCTTCAAGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGAC 540
DB 3965 GGCATCAAGTGAAGTCTTCAAGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGAC 4024
QY 541 GACCACTTCAAGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGAC 600
DB 4025 GACCACTTCAAGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGAC 4084
QY 601 TACCTGAGCAGCAGCTTCAAGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGAC 660
DB 4085 TACCTGAGCAGCAGCTTCAAGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGAC 4144
QY 661 CTGCTGAGTGTGAGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGAC 720
DB 4145 CTGCTGAGTGTGAGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGAC 4204
QY 721 CTTAAGCAGTGTGAGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGAC 780
DB 4205 CTTAAGCAGTGTGAGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGAC 4264
QY 781 TGTGCGCAGAGAGCGGAGTGAACGCTTCAAGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGAC 840
DB 4265 TGTGCGCAGAGAGCGGAGTGAACGCTTCAAGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGAC 4324
QY 841 GTGTAGATGC 850
DB 4325 GTGTAGATGC 4334

RESULT 3
US-10-811-081-98
; Sequence 98, Application US/10811081
; GENERAL INFORMATION:
; APPLICANT: El Tayar, Nabli

QY 121 GGCAAGCTGACCTTGAAGTTCATCTGCACACCGGACACCTCCGTCCTGCGCCACC 180
DB 3605 GGCAAGCTGACCTTGAAGTTCATCTGCACACCGGACACCTCCGTCCTGCGCCACC 3664
QY 181 CTCGTGACCACTTGAAGTTCATCTGCACACCGGACACCTCCGTCCTGCGCCACC 240
DB 3665 CTCGTGACCACTTGAAGTTCATCTGCACACCGGACACCTCCGTCCTGCGCCACC 3724
QY 241 GAGCAGCACTTCTTCAAGTTCGCGCAGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGAC 300
DB 3725 GAGCAGCACTTCTTCAAGTTCGCGCAGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGAC 3784
QY 301 TTCAAGAGCAGCAGCACTTCAAGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGAC 360
DB 3785 TTCAAGAGCAGCAGCACTTCAAGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGAC 3844
QY 361 GTGAACCGCATGAGCTGAAGGCGCATCTTCAAGGACGCGGACGCGGACGCGGACGCGGACGCGGAC 420
DB 3845 GTGAACCGCATGAGCTGAAGGCGCATCTTCAAGGACGCGGACGCGGACGCGGACGCGGACGCGGAC 3904
QY 421 AAGCTGAGTACCACTTCAAGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGAC 480
DB 3905 AAGCTGAGTACCACTTCAAGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGAC 3964
QY 481 GGCATCAAGTGAAGTCTTCAAGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGAC 540
DB 3965 GGCATCAAGTGAAGTCTTCAAGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGAC 4024
QY 541 GACCACTTCAAGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGAC 600
DB 4025 GACCACTTCAAGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGAC 4084
QY 601 TACCTGAGCAGCAGCTTCAAGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGAC 660
DB 4085 TACCTGAGCAGCAGCTTCAAGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGAC 4144
QY 661 CTGCTGAGTGTGAGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGAC 720
DB 4145 CTGCTGAGTGTGAGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGAC 4204
QY 721 CTTAAGCAGTGTGAGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGAC 780
DB 4205 CTTAAGCAGTGTGAGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGAC 4264
QY 781 TGTGCGCAGAGAGCGGAGTGAACGCTTCAAGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGAC 840
DB 4265 TGTGCGCAGAGAGCGGAGTGAACGCTTCAAGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGAC 4324
QY 841 GTGTAGATGC 850
DB 4325 GTGTAGATGC 4334

Query Match 84.6%; Score 719.2; DB 6; Length 14083;
Best Local Similarity 99.6%; Pred. No. 1.1e-125;
Matches 721; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

US-10-789-400-3
; PRIOR FILING DATE: 2003-06-13
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 14083
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HMPV strain 83 with GFP inserted prior to N gene.

QY 421 AAGCTGAGTACAACTACAAAGCCAAAGTCTATATCTATGCGCAGCAAGCAAGAAAC 480
 Db 2136 AAGCTGAGTACAACTACAAAGCCAAAGTCTATATCTATGCGCAGCAAGCAAGAAAC 2195
 QY 481 GGCATTAAGTGAACCTTCAAGATCCGCCCAACAATGAGAGAGCGGCTGACGCTGCC 540
 Db 2196 GGCATTAAGTGAACCTTCAAGATCCGCCCAACAATGAGAGAGCGGCTGACGCTGCC 2255
 QY 541 GACCACTACAGAGAAACCCCATCGGAGACGCGCCGCTGCTGCTGCCCAACAACAC 600
 Db 2256 GACCACTACAGAGAAACCCCATCGGAGACGCGCCGCTGCTGCTGCCCAACAACAC 2315
 QY 601 TACCTGAGCAACCAAGTCCGCCCTGAGCAAAAGACCCCAAGAGAGCGCATGACATGCTC 660
 Db 2316 TACCTGAGCAACCAAGTCCGCCCTGAGCAAAAGACCCCAAGAGAGCGCATGACATGCTC 2375
 QY 661 CTGCTGAGTGTGTGACCCGCCCGCGGATGATCTTGGCATGAGCAAGCTGTACAAAG 720
 Db 2376 CTGCTGAGTGTGTGACCCGCCCGCGGATGATCTTGGCATGAGCAAGCTGTACAAAG 2435
 QY 721 CTTAGCCATGGCT 733
 Db 2436 AGCGGCCCGGACT 2448

RESULT 5

us-10-826-523-52/c
 ; Sequence 52, Application US/10826523
 ; GENERAL INFORMATION:
 ; APPLICANT: FRASER JR., MALCOLM J.
 ; APPLICANT: LI, XU
 ; APPLICANT: BEAM, TERESA
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
 ; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
 ; TITLE OF INVENTION: VECTOR PIGGYBAC
 ; FILE REFERENCE: 835910-92098
 ; CURRENT APPLICATION NUMBER: US/10/826,523
 ; CURRENT FILING DATE: 2004-04-19
 ; PRIOR APPLICATION NUMBER: US/10/001,189
 ; PRIOR FILING DATE: 2001-10-30
 ; PRIOR APPLICATION NUMBER: 60/244,984
 ; PRIOR FILING DATE: 2000-11-01
 ; PRIOR APPLICATION NUMBER: 60/244,677
 ; PRIOR FILING DATE: 2000-10-31
 ; NUMBER OF SEQ ID NOS: 70
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 52
 ; LENGTH: 4952
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: pXL-Bac-EgFP
 ; OTHER INFORMATION: sequence
 us-10-826-523-52

Query Match 84.5%; Score 718.6; DB 6; Length 4952;
 Best Local Similarity 98.8%; Pred. No. 1,4e-125;
 Matches 724; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 ATGTGTAGCAAGGCGAGAGCTGTTCACCGGGGTGTGCTCCCTCTGTGTCAGCTGAC 60
 Db 2187 ATGTGTAGCAAGGCGAGAGCTGTTCACCGGGGTGTGCTCCCTCTGTGTCAGCTGAC 2128
 QY 61 GCGCAGTAAACGGCGACAAAGTTCAAGCGTCCGCGAGAGGCGAGGCGATGCACTTAC 120
 Db 2127 GCGCAGTAAACGGCGACAAAGTTCAAGCGTCCGCGAGAGGCGAGGCGATGCACTTAC 2068
 QY 121 GCGCAGTAAACGGCGACAAAGTTCAAGCGTCCGCGAGAGGCGAGGCGATGCACTTAC 180
 Db 2067 GCGCAGTAAACGGCGACAAAGTTCAAGCGTCCGCGAGAGGCGAGGCGATGCACTTAC 2008
 QY 181 CTGTGATCAACCTGATCTTACGAGGCGATGCTTCAAGCGCTAACCCCGACCATGAAG 240

Db 2007 CTGTGATCAACCTGATCTTACGAGGCGATGCTTCAAGCGCTAACCCCGACCATGAAG 1948
 QY 241 CAGCAGCACTTCTTCAAGTCCGCCATGCGCGAAGGCTACGTCAGAGAGCGACCATCTTC 300
 Db 1947 CAGCAGCACTTCTTCAAGTCCGCCATGCGCGAAGGCTACGTCAGAGAGCGACCATCTTC 1888
 QY 301 TTCAAGAGCAGAGGCACTTCAAGACCCCGCGCAAGGTGAAGTTTGAAGGCGACACCTTG 360
 Db 1887 TTCAAGAGCAGAGGCACTTCAAGACCCCGCGCAAGGTGAAGTTTGAAGGCGACACCTTG 1828
 QY 361 GTGAACCGCATGAGCTGAAGGCGATGCACTTCAAGAGAGAGCGGTAACATCTTGGGGAC 420
 Db 1827 GTGAACCGCATGAGCTGAAGGCGATGCACTTCAAGAGAGAGCGGTAACATCTTGGGGAC 1768
 QY 421 AAGCTGAGTCAACTACAAAGCCCAAGCGTATATCATGCGCGCAAGCAAGAAAC 480
 Db 1767 AAGCTGAGTCAACTACAAAGCCCAAGCGTATATCATGCGCGCAAGCAAGAAAC 1708
 QY 481 GGCATCAAGGTGAACCTTCAAGATCCGCCCAACAATGAGAGAGCGAGGCTGACGCTGCC 540
 Db 1707 GGCATCAAGGTGAACCTTCAAGATCCGCCCAACAATGAGAGAGCGAGGCTGACGCTGCC 1648
 QY 541 GACCACTACAGAGAAACCCCATCGGCGAAGGCGCGCGCGCTGCTGCTGCCCAACAAC 600
 Db 1647 GACCACTACAGAGAAACCCCATCGGCGAAGGCGCGCGCGCTGCTGCTGCCCAACAAC 1588
 QY 601 TACCTGAGCAACCAAGTCCGCCCTGAGCAAAAGACCCCAAGAGAGCGCATGACATGCTC 660
 Db 1587 TACCTGAGCAACCAAGTCCGCCCTGAGCAAAAGACCCCAAGAGAGCGCATGACATGCTC 1528
 QY 661 CTGCTGAGTGTGTGACCCGCCCGCGGATGATCTTGGCATGAGCAAGCTGTACAAAG 720
 Db 1527 CTGCTGAGTGTGTGACCCGCCCGCGGATGATCTTGGCATGAGCAAGCTGTACAAAG 1468
 QY 721 CTTAGCCATGGCT 733
 Db 1467 AGCGGCCCGGACT 1455

RESULT 6

us-10-485-281-6/c
 ; Sequence 6, Application US/10485281
 ; GENERAL INFORMATION:
 ; APPLICANT: Sandig, Volker
 ; APPLICANT: Jordan, Ingo
 ; TITLE OF INVENTION: RNA Amplification System Using Plant Components in
 ; TITLE OF INVENTION: Animal Cells
 ; FILE REFERENCE: 040026us/3H
 ; CURRENT APPLICATION NUMBER: US/10/485,281
 ; CURRENT FILING DATE: 2004-01-29
 ; PRIOR APPLICATION NUMBER: PCT/DE02/02863
 ; PRIOR FILING DATE: 2002-07-29
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 6
 ; LENGTH: 5451
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Vector pLMO-79
 us-10-485-281-6

Query Match 84.5%; Score 718.6; DB 6; Length 5451;
 Best Local Similarity 98.8%; Pred. No. 1,4e-125;
 Matches 724; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 ATGTGTAGCAAGGCGAGAGCTGTTCACCGGGGTGTGCTCCCTCTGTGTCAGCTGAC 60
 Db 4898 ATGTGTAGCAAGGCGAGAGCTGTTCACCGGGGTGTGCTCCCTCTGTGTCAGCTGAC 4839
 QY 61 GCGCAGTAAACGGCGACAAAGTTCAAGCGTCCGCGAGAGGCGAGGCGATGCACTTAC 120
 Db 4838 GCGCAGTAAACGGCGACAAAGTTCAAGCGTCCGCGAGAGGCGAGGCGATGCACTTAC 4779

Db 1585 TACCTGAGCACCAGTCCGCTGAGCAAAAGACCCCAAGAGCGGATCATGATC 1644
Qy 661 CTGCTGAGTTCTGTAACCGCCGCGGATCTCTCGGATGAGAGCTGTACAAAG 719
Db 1645 CTGCTGAGTTCTGTAACCGCCGCGGATCTCTCGGATGAGAGCTGTACAAAG 1703

RESULT 10

PCT-US04-06378-6
; Sequence 6, Application PC/TUS0406378
; GENERAL INFORMATION:
; APPLICANT: Avigenics, Inc
; TITLE OF INVENTION: Avian Integrase-mediated Transformation
; FILE REFERENCE: A181 1080.1
; CURRENT APPLICATION NUMBER: PCT/US04/06378
; CURRENT FILING DATE: 2004-03-05
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 5041
; TYPE: DNA
; ORGANISM: Plasmid pCMV-EGFP-attB
PCT-US04-06378-6

Query Match 84.4%; Score 717.4; DB 1; Length 5041;
Best Local Similarity 99.2%; Pred. No. 2,3e-125;
Matches 721; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 ATGTGAGCAAGGCGGAGAGCTTTCAACCGGGTGTGCTCCATCTCTGTGAGCTGGAC 60
Db 4301 ATGTGAGCAAGGCGGAGAGCTTTCAACCGGGTGTGCTCCATCTCTGTGAGCTGGAC 4360
Qy 61 GGGCAGCTAAACGGCCCAAGTTGAGGCTGTCCGCGCAAGGCGGAGGCGATGCCACTAC 120
Db 4361 GGGCAGCTAAACGGCCCAAGTTGAGGCTGTCCGCGCAAGGCGGAGGCGATGCCACTAC 4420
Qy 121 GGGCAGCTGACCTCTGAAGTTCTATGCAACACCGGCAAGCTGCGCTGCTGCGCCAC 180
Db 4421 GGGCAGCTGACCTCTGAAGTTCTATGCAACACCGGCAAGCTGCGCTGCTGCGCCAC 4480
Qy 181 CTGTGACCACTCTGACCTTCAAGCTGTGCAAGCTTCAAGCTTCAAGCTTCAAGCTT 240
Db 4481 CTGTGACCACTCTGACCTTCAAGCTGTGCAAGCTTCAAGCTTCAAGCTTCAAGCTT 4540
Qy 241 CAGCAGCACTCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTT 300
Db 4541 CAGCAGCACTCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTT 4600
Qy 301 TTCAAGGACGACGCACTCAAGCAACCGCGCGGAGTGAAGTTCAGAGGCGACACTTC 360
Db 4601 TTCAAGGACGACGCACTCAAGCAACCGCGCGGAGTGAAGTTCAGAGGCGACACTTC 4660
Qy 361 GTGAACCGCATCGAGCTGAAGGCGATCTTCAAGAGGAGCGGCAATCTTGGGGAC 420
Db 4661 GTGAACCGCATCGAGCTGAAGGCGATCTTCAAGAGGAGCGGCAATCTTGGGGAC 4720
Qy 421 AAGCTGAGTCACTCAAGCAACGCTCTATCATGAGCGCAACAGCAAGAGAC 480
Db 4721 AAGCTGAGTCACTCAAGCAACGCTCTATCATGAGCGCAACAGCAAGAGAGAC 4780
Qy 481 GGCATCAAGTGAATCTTCAAGTCCGCAACATCGAGAGCGGAGGCTGAGCTGCC 540
Db 4781 GGCATCAAGTGAATCTTCAAGTCCGCAACATCGAGAGCGGAGGCTGAGCTGCC 4840
Qy 541 GACCACTACGACGAGAAACCCCATGCGGAGCGGCGCTGCTGCTGCTGCTGCTGCTG 600
Db 4841 GACCACTACGACGAGAAACCCCATGCGGAGCGGCGCTGCTGCTGCTGCTGCTGCTG 4900
Qy 601 TACCTGAGCAACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Db 4901 TACCTGAGCAACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4960
Qy 661 CTGCTGAGTTCTGTAACCGCCGCGGATCTCTCGGATGAGAGCTGTACAAAG 720

Db 4961 CTGCTGAGTTCTGTAACCGCCGCGGATCTCTCGGATGAGAGCTGTACAAAG 5020
Qy 721 CTTAAGC 727
Db 5021 AGCGGCC 5027

RESULT 11

US-10-790-455-6
; Sequence 6, Application US/10790455
; GENERAL INFORMATION:
; APPLICANT: Avigenics, Inc
; TITLE OF INVENTION: Avian Integrase-mediated Transformation
; FILE REFERENCE: A181 1080.1
; CURRENT APPLICATION NUMBER: US/10/790,455
; CURRENT FILING DATE: 2004-03-01
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 5041
; TYPE: DNA
; ORGANISM: Plasmid pCMV-EGFP-attB
US-10-790-455-6

Query Match 84.4%; Score 717.4; DB 6; Length 5041;
Best Local Similarity 99.2%; Pred. No. 2,3e-125;
Matches 721; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 ATGTGAGCAAGGCGGAGAGCTTTCAACCGGGTGTGCTCCATCTCTGTGAGCTGGAC 60
Db 4301 ATGTGAGCAAGGCGGAGAGCTTTCAACCGGGTGTGCTCCATCTCTGTGAGCTGGAC 4360
Qy 61 GGGCAGCTAAACGGCCCAAGTTGAGGCTGTCCGCGCAAGGCGGAGGCGATGCCACTAC 120
Db 4361 GGGCAGCTAAACGGCCCAAGTTGAGGCTGTCCGCGCAAGGCGGAGGCGATGCCACTAC 4420
Qy 121 GGGCAGCTGACCTCTGAAGTTCTATGCAACACCGGCAAGCTGCGCTGCTGCGCCAC 180
Db 4421 GGGCAGCTGACCTCTGAAGTTCTATGCAACACCGGCAAGCTGCGCTGCTGCGCCAC 4480
Qy 181 CTGTGACCACTCTGACCTTCAAGCTGTGCAAGCTTCAAGCTTCAAGCTTCAAGCTT 240
Db 4481 CTGTGACCACTCTGACCTTCAAGCTGTGCAAGCTTCAAGCTTCAAGCTTCAAGCTT 4540
Qy 241 CAGCAGCACTCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTT 300
Db 4541 CAGCAGCACTCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTT 4600
Qy 301 TTCAAGGACGACGCACTCAAGCAACCGCGCGGAGTGAAGTTCAGAGGCGACACTTC 360
Db 4601 TTCAAGGACGACGCACTCAAGCAACCGCGCGGAGTGAAGTTCAGAGGCGACACTTC 4660
Qy 361 GTGAACCGCATCGAGCTGAAGGCGATCTTCAAGAGGAGCGGCAATCTTGGGGAC 420
Db 4661 GTGAACCGCATCGAGCTGAAGGCGATCTTCAAGAGGAGCGGCAATCTTGGGGAC 4720
Qy 421 AAGCTGAGTCACTCAAGCAACGCTCTATCATGAGCGCAACAGCAAGAGAC 480
Db 4721 AAGCTGAGTCACTCAAGCAACGCTCTATCATGAGCGCAACAGCAAGAGAGAC 4780
Qy 481 GGCATCAAGTGAATCTTCAAGTCCGCAACATCGAGAGCGGAGGCTGAGCTGCC 540
Db 4781 GGCATCAAGTGAATCTTCAAGTCCGCAACATCGAGAGCGGAGGCTGAGCTGCC 4840
Qy 541 GACCACTACGACGAGAAACCCCATGCGGAGCGGCGCTGCTGCTGCTGCTGCTGCTG 600
Db 4841 GACCACTACGACGAGAAACCCCATGCGGAGCGGCGCTGCTGCTGCTGCTGCTGCTG 4900
Qy 601 TACCTGAGCAACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Db 4901 TACCTGAGCAACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4960

QY 661 CTGCTGAGTTGTCAGCCGCGGATCACTCTGCGATGAGAGAGCTGACAGAG 720
 Db 4961 CTGCTGAGTTGTCAGCCGCGGATCACTCTGCGATGAGAGAGCTGACAGAG 5020
 QY 721 CTTAGCC 727
 Db 5021 AGCGGCC 5027
 RESULT 12
 US-10-475-962-54
 ; Sequence 54, Application US/10475962
 ; GENERAL INFORMATION:
 ; APPLICANT: CHAMON, PIERRE
 ; APPLICANT: GHISELINCK, NORBERT B.
 ; APPLICANT: SCHNUTGEN, FRANK
 ; TITLE OF INVENTION: METHOD FOR THE STABLE INVERSION OF DNA SEQUENCE BY SITE-SPECIFIC
 ; FILE REFERENCE: 37991-0025
 ; CURRENT APPLICATION NUMBER: US/10/475,962
 ; PRIOR FILING DATE: 2003-10-27
 ; PRIOR APPLICATION NUMBER: PCT/IB02/02493
 ; PRIOR FILING DATE: 2002-04-19
 ; PRIOR APPLICATION NUMBER: US 09/843,150
 ; NUMBER OF SEQ ID NOS: 57
 ; SOFTWARE: PatentIn Ver. 3.2
 ; SEQ ID NO 54
 ; LENGTH: 8693
 ; TYPE: DNA
 ; ORGANISM: Artificial sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial sequence: DNA
 ; OTHER INFORMATION: sequence of plasmid pLEXR
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1)..(360)
 ; OTHER INFORMATION: SV40 promoter, sense
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1130)
 ; OTHER INFORMATION: loxP1 site, sense
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1050)
 ; OTHER INFORMATION: loxP1 site, sense
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1170)..(2050)
 ; OTHER INFORMATION: EGFP polyA gene, sense
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (2060)..(5700)
 ; OTHER INFORMATION: lacZ polyA gene, antisense
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (5710)
 ; OTHER INFORMATION: loxP1 site, antisense
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (5790)
 ; OTHER INFORMATION: lox511 site, antisense
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (5830)..(8693)
 ; OTHER INFORMATION: vector sequence
 ; US-10-475-962-54
 Query Match 84.4%, Score 717.4, DB 6, Length 8693;

Best Local Similarity 99.9%; Pred. No. 2.4e-125;
 Matches 718; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ATGCTGAGCAAGGCGGAGAGCTGTCACCGGAGGTGGTGGCCATCTGCTGAGCTGAG 60
 Db 1172 ATGCTGAGCAAGGCGGAGAGCTGTCACCGGAGGTGGTGGCCATCTGCTGAGCTGAG 1231
 QY 61 GGGCAGTAAACGGCCCAAGTTAGCGTGTCCGGCGAGAGGCGAGGCGATGCCACTTAC 120
 Db 1232 GGGCAGTAAACGGCCCAAGTTAGCGTGTCCGGCGAGAGGCGAGGCGATGCCACTTAC 1291
 QY 121 GGGCAGTAAACGGCCCAAGTTAGCGTGTCCGGCGAGAGGCGAGGCGATGCCACTTAC 180
 Db 1292 GGGCAGTAAACGGCCCAAGTTAGCGTGTCCGGCGAGAGGCGAGGCGATGCCACTTAC 1351
 QY 181 CTGCTGAGCAAGGCGGAGAGCTGTCACCGGAGGTGGTGGCCATCTGCTGAGCTGAG 240
 Db 1352 CTGCTGAGCAAGGCGGAGAGCTGTCACCGGAGGTGGTGGCCATCTGCTGAGCTGAG 1411
 QY 241 CAGCAGCACTTCTTCAAGTCCGCCATGCCGAGAGGCTACGTCAGAGCGCAGCATCTTC 300
 Db 1412 CAGCAGCACTTCTTCAAGTCCGCCATGCCGAGAGGCTACGTCAGAGCGCAGCATCTTC 1471
 QY 301 TTCAAGAGCAAGGCGGAGAGCTGTCACCGGAGGTGGTGGCCATCTGCTGAGCTGAG 360
 Db 1472 TTCAAGAGCAAGGCGGAGAGCTGTCACCGGAGGTGGTGGCCATCTGCTGAGCTGAG 1531
 QY 361 GTGAGCCGATTCGAGTGAAGGAGGAGTGAAGTTCAGAGAGAGGCGCAATCTGCTGAG 420
 Db 1532 GTGAGCCGATTCGAGTGAAGGAGGAGTGAAGTTCAGAGAGAGGCGCAATCTGCTGAG 1531
 QY 421 AAGCTGAGTAACTTCAACAGGCAAGGCGTATATCATATGCGGCAAGAGAGAGAG 480
 Db 1592 AAGCTGAGTAACTTCAACAGGCAAGGCGTATATCATATGCGGCAAGAGAGAGAG 1651
 QY 481 GGCATCAAGGTGAAGTTCAGATTCGCGCAACATCATGAGAGAGGCGACGTCAGCTGCC 540
 Db 1652 GGCATCAAGGTGAAGTTCAGATTCGCGCAACATCATGAGAGAGGCGACGTCAGCTGCC 1711
 QY 541 GACCACTACAGGAGAGAGAGGCGGAGGAGGCGGAGGCGGAGGCGGAGGCGGAGG 600
 Db 1712 GACCACTACAGGAGAGAGAGGCGGAGGAGGCGGAGGCGGAGGCGGAGGCGGAGG 1771
 QY 601 TACCTGAGCAAGGCGGAGGAGGCGGAGGAGGCGGAGGCGGAGGCGGAGGCGGAGG 660
 Db 1772 TACCTGAGCAAGGCGGAGGAGGCGGAGGAGGCGGAGGCGGAGGCGGAGGCGGAGG 1831
 QY 661 CTGCTGAGTTGTCAGCCGCGGATCACTCTGCGATGAGAGAGCTGACAGAG 719
 Db 1832 CTGCTGAGTTGTCAGCCGCGGATCACTCTGCGATGAGAGAGCTGACAGAG 1890
 RESULT 13
 US-10-785-862-9
 ; Sequence 9, Application US/10785862
 ; GENERAL INFORMATION:
 ; APPLICANT: Donald Danforth Plant Science Center
 ; APPLICANT: Fauquet, Claude M.
 ; APPLICANT: Padmanabhan, Chellappan
 ; APPLICANT: Ramachandran, Vanitharani
 ; TITLE OF INVENTION: siRNA-mediated inhibition of gene expression in plant cells
 ; FILE REFERENCE: C35621/104850
 ; CURRENT APPLICATION NUMBER: US/10/785,862
 ; PRIOR FILING DATE: 2004-02-24
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 9
 ; LENGTH: 717
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: GFP coding sequence
 ; US-10-785-862-9

OTHER INFORMATION: n = A,T,C or G
US-10-672-764A-53

Query Match 84.4%; Score 717; DB 6; Length 10089;
Best Local Similarity 100.0%; Pred. No. 2.9e-125;
Matches 717; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGCTGACGAAAGGCGAGAGCTTTACCGGGGTGTGCTCCATCTGTGCTGAGCTGAGC	60
DB	1029	ATGCTGACGAAAGGCGAGAGCTTTACCGGGGTGTGCTCCATCTGTGCTGAGCTGAGC	1088
QY	61	GGCGACGTAAAGCGCCACAAAGTTGACGGTCCGGGCGAGGGGAGGGGATGCCCTTAC	120
DB	1089	GGCGACGTAAAGCGCCACAAAGTTGACGGTCCGGGCGAGGGGAGGGGATGCCCTTAC	1148
QY	121	GGCAAGCTGACCTTGAAGTTCAATCGACCAACCGGCAAGCTGCCGTGCCCTGAGCCAC	180
DB	1149	GGCAAGCTGACCTTGAAGTTCAATCGACCAACCGGCAAGCTGCCGTGCCCTGAGCCAC	1208
QY	181	CTCGTGAACACCTTACCTTACGAGGTGCTTCAAGCCGCTACCCGACCAATGAAG	240
DB	1209	CTCGTGAACACCTTACCTTACGAGGTGCTTCAAGCCGCTACCCGACCAATGAAG	1268
QY	241	CAGCAGCACTTCTTCAAGTCCGCCATGCGCCGAAAGCTAGTCCAGGAGGCAACATCTTC	300
DB	1269	CAGCAGCACTTCTTCAAGTCCGCCATGCGCCGAAAGCTAGTCCAGGAGGCAACATCTTC	1328
QY	301	TTCAAGGACGAGCGCACTTCAAGACCGCGCGCGAGGTGAAGTTGAGGGGCAACCTTG	360
DB	1329	TTCAAGGACGAGCGCACTTCAAGACCGCGCGCGAGGTGAAGTTGAGGGGCAACCTTG	1388
QY	361	GTGAACCGCATCGAGCTGAAAGGATCGACTTTCAAGGAGGACGGCAACATCTTGAGGAC	420
DB	1389	GTGAACCGCATCGAGCTGAAAGGATCGACTTTCAAGGAGGACGGCAACATCTTGAGGAC	1448
QY	421	AAGCTGAGTAACTAACCAAGCAAGCAAGCTTATATCATGAGCGCAACAGCAAGAAC	480
DB	1449	AAGCTGAGTAACTAACCAAGCAAGCAAGCTTATATCATGAGCGCAACAGCAAGAAC	1508
QY	481	GGCATCAAGGTGAAGCTTCAAGATCCGCCCAACATCGAGAGCGGCAAGCTGAGC	540
DB	1509	GGCATCAAGGTGAAGCTTCAAGATCCGCCCAACATCGAGAGCGGCAAGCTGAGC	1568
QY	541	GACCACTAACGAGGAAACACCCCGATCGGCGACGGCCCGCTGCTGCTGCCGACAC	600
DB	1569	GACCACTAACGAGGAAACACCCCGATCGGCGACGGCCCGCTGCTGCTGCCGACAC	1628
QY	601	TACCTGAGCAACCGATCGCCCTGAGCAAGACCCCAACGAGAGCGGATCACATGATC	660
DB	1629	TACCTGAGCAACCGATCGCCCTGAGCAAGACCCCAACGAGAGCGGATCACATGATC	1688
QY	661	CTGCTGAGATTGTCGACCGCGCGGAGTCACTCTGCGCATGACGAGCTGTACAAG	717
DB	1689	CTGCTGAGATTGTCGACCGCGCGGAGTCACTCTGCGCATGACGAGCTGTACAAG	1745

Search completed: May 15, 2004, 06:24:44
Job time : 173 secs